

RANDOM WALKS ON WREATH PRODUCTS OF GROUPS

BY CLYDE H. SCHOOLFIELD, JR.

Harvard University

We bound the rate of convergence to uniformity for certain random walks on the complete monomial groups $G \wr S_n$ for any group G . These results provide rates of convergence for random walks on a number of groups of interest: the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, the generalized symmetric group $\mathbb{Z}_m \wr S_n$, and $S_m \wr S_n$. These results provide benchmarks to which many other random walks, modeling a wide range of phenomena, may be compared using the comparison technique, thereby yielding bounds on the rates of convergence to uniformity for previously intractable random walks.

1. Introduction. How many steps does it take for a deck of n cards to achieve near-randomness if, at each step, two randomly chosen cards are transposed? This question was answered by Diaconis and Shahshahani (1981). Now suppose that at each step, the two cards are not only transposed, but also possibly flipped. Suppose that, instead of n cards, there are n wheels and that, at each step, two wheels are transposed and then spun. Or suppose that, instead of n wheels, there are n decks of cards and that, at each step, two decks are transposed and then shuffled. How many steps does it take for these processes to achieve near-randomness?

Or perhaps we are interested in a process that is (at least somewhat) similar in form to one of those described above. Can we determine how many steps are needed for it to achieve near-randomness? These are the types of questions that we consider.

For a certain random walk on the symmetric group S_n that is generated by random transpositions, Diaconis and Shahshahani (1981) obtained bounds on the rate of convergence to uniformity using group representation theory. Similarly, we bound the rate of convergence to uniformity for a random walk on the complete monomial group $G \wr S_n$ that is generated by random transpositions, followed by independent randomizations of the transposed elements. Specifically, we determine that $\frac{1}{2}n \log n + \frac{1}{4}n \log(|G| - 1)$ steps are both necessary and sufficient for ℓ^2 distance to become small. We also determine that $\frac{1}{2}n \log n$ steps are both necessary and sufficient for total variation distance to become small. These results provide

AMS 1991 subject classifications. Primary 60B15, 60J15; secondary 20E22.

Key words and phrases. Random walk, Markov chain, wreath product, group, Fourier transform, eigenvalue, comparison technique.

rates of convergence for random walks on a number of groups of interest: the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, the generalized symmetric group $\mathbb{Z}_m \wr S_n$, and $S_m \wr S_n$. In the special case of the hyperoctahedral group, the two rates of convergence are the same in both metrics. We also examine a slight variant of this random walk, establishing upper and lower bounds on its rate of convergence to uniformity.

The comparison technique was introduced by Diaconis and Saloff-Coste (1993) as a method for bounding the rate of convergence to uniformity of a symmetric random walk on a finite group by comparing it to a benchmark random walk whose rate of convergence is known. Our random walks on the hyperoctahedral, generalized symmetric, and complete monomial groups provide such benchmarks to which many other random walks, modeling a wide range of phenomena, may be compared, thereby yielding bounds on the rates of convergence to uniformity for previously intractable random walks. Schoolfield (1998) used the results of this paper to analyze two specific examples using this technique, one of which has applications in mathematical biology. Schoolfield (1998) further specialized the comparison technique to random walks generated by random transpositions along the edges of a graph and analyzed several examples.

Section 2 presents the basic properties and results from group theory and representation theory necessary to analyze random walks on groups. Section 3 extends the idea of random transpositions of n cards, analyzed in Section 2, to a set of n decks of m cards each and beyond.

2. Groups, Representations, and Random Walks.

2.1. Introduction. Imagine n cards, labeled 1 through n , on a table in sequential order. Independently choose two integers p and q uniformly from $\{1, 2, \dots, n\}$. If $p \neq q$, transpose the cards in positions p and q ; we denote this by $\tau \in S_n$, where τ is the transposition $(p\ q)$. We shall refer to this procedure as a *random transposition* of n cards. If $p = q$ (which occurs with probability $1/n$), leave the cards in their current positions. This action is of course the identity permutation, which is denoted by $e \in S_n$.

If this process is repeated many times, the cards will appear to be in uniformly random order, that is, will appear to be the result of a *random permutation*. This process may be modeled formally using a probability measure P (which we shall regard as a probability mass

function) on the symmetric group S_n , namely,

$$\begin{aligned} P(e) &:= \frac{1}{n} \text{ where } e \text{ is the identity element,} \\ P(\tau) &:= \frac{2}{n^2} \text{ where } \tau \text{ is any transposition,} \\ P(\pi) &:= 0 \text{ otherwise.} \end{aligned} \tag{2.1.1}$$

Repeating the process above k times is modeled as the convolution P^{*k} of the measure P with itself k times. Since there are $n!$ elements in S_n , the uniform probability measure on the set of all permutations of S_n is given by

$$U(\pi) := \frac{1}{n!} \text{ for every } \pi \in S_n. \tag{2.1.2}$$

The following result, which is Theorem 1 of Diaconis and Shahshahani (1981) and was later included in Diaconis (1988) as Theorem 5 in Section D of Chapter 3, establishes an upper bound on both the total variation distance and the ℓ^2 distance (both defined in Section 2.6) between P^{*k} and U .

THEOREM 2.1.3. *Let P and U be the probability measures on the symmetric group S_n defined in (2.1.1) and (2.1.2), respectively. Let $k = \frac{1}{2}n \log n + cn$. Then there exists a universal constant $a > 0$ such that*

$$\|P^{*k} - U\|_{\text{TV}} \leq \frac{1}{2} (n!)^{1/2} \|P^{*k} - U\|_2 \leq ae^{-2c} \text{ for all } c > 0.$$

In the following sections we present the results that were needed to prove this theorem and which are used to prove analogous results in later sections. In Section 2.2 we present basic definitions from group theory. In Section 2.3 we present basic results from the theory of group representations, while Section 2.4 concentrates on the characters of these representations. In Section 2.5 we introduce the Fourier transform, and we show how it may be used to bound the distance to uniformity of a random walk on a group in Section 2.6. In Section 2.7 we show how the results from these previous sections were applied to the random walk on the symmetric group defined above, proving Theorem 2.1.3 and a matching lower bound.

2.2. Group Theory. We now present basic definitions from the theory of groups which will be needed to conduct our analysis. A more detailed introduction to this subject may be found in Chapter 1 of Alperin and Bell (1995).

A *group* is a non-empty set G with a binary operation on G , usually called multiplication, which satisfies the following three axioms: (i) Multiplication is associative, (ii) There is a unique *identity* element $e \in G$, and (iii) For every $g \in G$ there is a unique *inverse* element

$g^{-1} \in G$ such that $gg^{-1} = e = g^{-1}g$. The number of elements of a finite group G is called the *order* of G and is denoted by $|G|$.

A subset H of a group G is called a *subgroup* of G if it forms a group under the multiplication of G restricted to H . A subgroup N of G is called a *normal subgroup* of G if $gNg^{-1} \subseteq N$ for all $g \in G$, where $gNg^{-1} := \{gng^{-1} \in G : n \in N\}$. If $gh = hg$ for all $g, h \in G$, then G is called an *abelian* group. Notice that every subgroup of an abelian group is a normal subgroup.

The set $kH := \{kh \in G : h \in H\}$, where $k \in G$ and H is a subgroup of G , is called a *left coset* of H in G . (Right cosets are defined analogously.) The set of all left cosets of H in G is called the *left coset space* and is denoted by G/H . If $|G/H| = n$, a set of *left coset representatives* $\{k_1, k_2, \dots, k_n\}$ may be chosen so that the left cosets k_1H, k_2H, \dots, k_nH comprise precisely the coset space G/H . If N is a normal subgroup of G , then G/N is also a group known as the *quotient group*.

One group of interest to us is the *cyclic group* \mathbb{Z}_n , which is the set $\{0, 1, 2, \dots, n-1\}$ under the operation of addition mod n . Notice that \mathbb{Z}_n is abelian and that $|\mathbb{Z}_n| = n$. Another group of interest to us is the *symmetric group* S_n , which is the set of all permutations of $\{1, 2, \dots, n\}$ under the operation of composition of functions (where we compose functions from right to left). Notice that S_n is *not* abelian for any $n \geq 3$, and that $|S_n| = n!$.

Two elements g and h of G are called *conjugate* if there exists some $k \in G$ such that $h = kgk^{-1}$. The set of elements conjugate to a particular element $g \in G$ form the *conjugacy class* of g . Conjugacy is an equivalence relation and partitions a group G into disjoint subsets, namely, the conjugacy classes.

The set of ordered pairs of elements of groups G_1 and G_2 with multiplication defined componentwise is called the *direct product* of G_1 and G_2 and is denoted by $G_1 \times G_2$. The group of elements $(g_1, g_2, \dots, g_n; \pi) \in G^n \times S_n$, with multiplication defined by

$$(h_1, \dots, h_n; \sigma) \cdot (g_1, \dots, g_n; \pi) = (h_1 \cdot g_{\sigma^{-1}(1)}, \dots, h_n \cdot g_{\sigma^{-1}(n)}; \sigma\pi),$$

is called the *wreath product* of G with S_n and is denoted $G \wr S_n$. The identity element of $G \wr S_n$ is $(e, \dots, e; e)$, and in $G \wr S_n$ we have $(g_1, \dots, g_n; \pi)^{-1} = (g_{\pi(1)}^{-1}, \dots, g_{\pi(n)}^{-1}; \pi^{-1})$. This definition is equivalent to saying that $G \wr S_n$ is the *semidirect product* of G^n with S_n , where the action of S_n on G^n is $\pi(g_1, g_2, \dots, g_n) = (g_{\pi^{-1}(1)}, g_{\pi^{-1}(2)}, \dots, g_{\pi^{-1}(n)})$. For a general definition of semidirect product see Alperin and Bell (1995) or Simon (1996). Three special cases of particular interest to us are the *hyperoctahedral group* $\mathbb{Z}_2 \wr S_n$, the *generalized symmetric group* $\mathbb{Z}_m \wr S_n$, and $S_m \wr S_n$.

2.3. Representation Theory. We now present basic properties and results from the representation theory of groups which will be needed to conduct our analysis. A more detailed introduction to this subject may be found in Chapters 1 through 3 of Serre (1977) or in Chapter 2 of Diaconis (1988). Other sources include Alperin and Bell (1995) and Simon (1996).

Suppose that V is a finite-dimensional vector space over the complex numbers. The *general linear group* $\mathbf{GL}(V)$ is the group of isomorphisms of V onto itself. An element of $\mathbf{GL}(V)$ is a linear mapping of V onto V . Such a map has an inverse which is also linear.

A *representation* of a finite group G in V is a homomorphism $\rho : G \longrightarrow \mathbf{GL}(V)$. The choice of a basis for V assigns an invertible matrix $\rho(g)$ to each $g \in G$. There exists $\rho(g) \in \mathbf{GL}(V)$ for each $g \in G$ such that $\rho(g_1g_2) = \rho(g_1)\rho(g_2)$ for all $g_1, g_2 \in G$. This implies that $\rho(e) = I$ and that $\rho(g^{-1}) = \rho(g)^{-1}$.

Two representations of G , say ρ_1 in V_1 and ρ_2 in V_2 , are said to be *isomorphic* (or *equivalent*) if there exists a linear isomorphism $\mu : V_1 \longrightarrow V_2$ such that $\mu \circ \rho_1(g) = \rho_2(g) \circ \mu$ for all $g \in G$. The *dimension* (or *degree*) of ρ is defined to be the dimension of V and is denoted by d_ρ . Since we shall be interested in representations only up to equivalence, we may without loss of generality assume that $V = \mathbb{C}^n$ when $d_\rho = n$. The *trivial representation* is the one-dimensional representation with $V = \mathbb{C}$ that sends every element of G to 1.

A vector subspace $W \subseteq V$ is *stable* under G if $\rho(g)w \in W$ for all $w \in W$ and all $g \in G$. If $W \subseteq V$ is stable under G , then ρ restricted to $\mathbf{GL}(W)$ gives a *subrepresentation* of ρ . A representation ρ is *irreducible* if the only vector subspaces of V which are stable under G are V itself and the trivial subspace. An important relationship between irreducible representations and conjugacy classes is given by the following, which is Theorem 7 in Section 2.5 of Serre (1977).

PROPOSITION 2.3.1. *The number of nonisomorphic irreducible representations of G equals the number of conjugacy classes of G .*

An important relationship between the dimensions of the irreducible representations of G and the order of G is given by the following, which is Corollary 2(a) of Proposition 5 in Section 2.4 of Serre (1977).

LEMMA 2.3.2. *Suppose that $\{\rho_1, \rho_2, \dots, \rho_s\}$ is a complete set of nonisomorphic irreducible representations of G with dimensions $d_{\rho_1}, d_{\rho_2}, \dots, d_{\rho_s}$, respectively. Then*

$$\sum_{i=1}^s d_{\rho_i}^2 = |G|.$$

It follows directly from Proposition 2.3.1 and Lemma 2.3.2 that a group G is abelian if and only if all of its irreducible representations are one-dimensional.

The following useful result is *Schur's Lemma*, which is Proposition 4 in Section 2.2 of Serre (1977).

LEMMA 2.3.3. *Suppose that $\rho_1 : G \longrightarrow \mathbf{GL}(V_1)$ and $\rho_2 : G \longrightarrow \mathbf{GL}(V_2)$ are irreducible representations of G and that $\mu : V_1 \longrightarrow V_2$ is such that $\mu \circ \rho_1(g) = \rho_2(g) \circ \mu$ for all $g \in G$. Then (i) if ρ_1 and ρ_2 are not isomorphic, it follows that $\mu = 0$; and (ii) if $V_1 = V_2$ and $\rho_1 = \rho_2$, it follows that μ is a constant times the identity.*

The *direct sum* $A \oplus B$ of an $m \times m$ matrix A and an $n \times n$ matrix B is the $(m+n) \times (m+n)$ block diagonal matrix $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. The direct sum $\rho_1 \oplus \rho_2$ of two representations is then defined by $(\rho_1 \oplus \rho_2)(g) = \rho_1(g) \oplus \rho_2(g)$. By use of the direct sum, the irreducible representations of G can be used to construct all other representations of G , as described in the following, which is Theorem 2 in Section 1.4 of Serre (1977).

PROPOSITION 2.3.4. *Every representation of G is the direct sum of irreducible representations of G .*

The *tensor product* $A \otimes B$ of an $m \times m$ matrix A and an $n \times n$ matrix B is an $mn \times mn$ matrix which is constructed in the following manner. Begin with an $m \times m$ block matrix in which each of the m^2 blocks is the matrix B . Then multiply the block in position (i, j) by the scalar $a_{ij} \in A$ for $1 \leq i, j \leq m$. The tensor product $\rho_1 \otimes \rho_2$ of two representations is then defined by $(\rho_1 \otimes \rho_2)(g) = \rho_1(g) \otimes \rho_2(g)$. By use of the tensor product, the irreducible representations of G_1 and G_2 can be used to construct all the irreducible representations of their direct product, as described in the following, which is Theorem 10 in Section 3.2 of Serre (1977).

PROPOSITION 2.3.5. *Suppose that ρ_1 and ρ_2 are irreducible representations of G_1 and G_2 , respectively. Then $\rho_1 \otimes \rho_2$ is an irreducible representation of $G_1 \times G_2$. Furthermore, each irreducible representation of $G_1 \times G_2$ is isomorphic to such a representation.*

Suppose that H is a subgroup of G and that ρ is a representation of G . A representation $\rho \downarrow_H^G$ of H , known as the *restricted representation*, can be constructed from ρ by defining

$$\rho \downarrow_H^G (h) := \rho(h) \quad \text{for each } h \in H.$$

Now suppose that H is a subgroup of G and that ρ is a representation of H . Also suppose that $\{k_1, k_2, \dots, k_n\}$ is a complete set of left coset representatives of H in G . A representation $\rho \uparrow_H^G$ of G , known as the *induced representation*, can be constructed from ρ by defining, for each $g \in G$, an $n \times n$ block matrix $\rho \uparrow_H^G (g)$ whose block in position (i, j) , for $1 \leq i, j \leq n$, is the $d_\rho \times d_\rho$ matrix

$$\rho \uparrow_H^G (g)_{ij} := \begin{cases} \rho(k_i^{-1} g k_j) & \text{if } k_i^{-1} g k_j \in H, \\ 0 & \text{if } k_i^{-1} g k_j \notin H. \end{cases}$$

The induced representation does not depend on the choice of coset representatives.

2.4. Character Theory. The *character* of the representation ρ at the element $g \in G$ is defined to be the trace of $\rho(g)$ and is denoted by $\chi_\rho(g)$. Notice that the character of a representation is independent of the choice of basis of V . The characters of the irreducible representations of G are called *irreducible characters*. The choice of the term “character” is to emphasize that it characterizes the representation; according to Corollary 2 of Theorem 4 in Section 2.3 of Serre (1977), two representations are isomorphic if and only if they have the same character.

Some important properties of characters are given in the following, which is Proposition 1 in Section 2.1 of Serre (1977).

LEMMA 2.4.1. *Suppose that χ is the character of a representation ρ of G with dimension d_ρ . Then*

- (a) $\chi(e) = d_\rho$ for $e \in G$,
- (b) $\chi(g^{-1}) = \overline{\chi(g)}$ for $g \in G$,
- (c) $\chi(hgh^{-1}) = \chi(g)$ for $g, h \in G$.

The characters of direct sums and tensor products may be easily calculated by use of the following, which is Proposition 2 in Section 2.1 of Serre (1977).

LEMMA 2.4.2. *Suppose that ρ_1 and ρ_2 are representations of G with characters χ_1 and χ_2 , respectively. Then the character of the direct sum $\rho_1 \oplus \rho_2$ is $\chi_1 + \chi_2$ and the character of the tensor product $\rho_1 \otimes \rho_2$ is $\chi_1 \cdot \chi_2$.*

The characters of induced representations may be calculated by use of the following, which is Corollary 6 in Section 16 of Alperin and Bell (1995).

LEMMA 2.4.3. *Suppose that χ_H is the character of an irreducible representation ρ_H of a subgroup H of a finite group G . For $g \in G$, suppose that the number t of conjugacy classes of H whose members are conjugate in G to g is positive. Let h_1, h_2, \dots, h_t be representatives of these t conjugacy classes of H and let k_1, k_2, \dots, k_t be the sizes of these classes. Let ℓ be the size of the conjugacy class of g in G . Then the value at g of the character χ of the induced representation $\rho_H \uparrow_H^G$ of G is given by*

$$\chi(g) = \frac{|G|}{|H|} \sum_{i=1}^t \frac{k_i}{\ell} \chi_H(h_i).$$

Suppose that f_1 and f_2 are functions on a group G . The *inner product* of f_1 and f_2 is defined by

$$\langle f_1, f_2 \rangle_G := \frac{1}{|G|} \sum_{g \in G} f_1(g) \overline{f_2(g)}.$$

A function on G which is constant on each conjugacy class of G is called a *class function*. Notice that Lemma 2.4.1(c) asserts that the character χ of a representation ρ of G is a class function. An important relationship between the irreducible characters of G and the space of class functions on G is given by the following, which is Theorem 6 in Section 2.5 of Serre (1977).

PROPOSITION 2.4.4. *The characters $\chi_1, \chi_2, \dots, \chi_s$ of the irreducible representations of G form an orthonormal basis for the Hilbert space of class functions on G with respect to the inner product $\langle \cdot, \cdot \rangle$ defined above.*

2.5. The Fourier Transform. We now introduce an extremely important tool for our analysis. Suppose that P is a function on G . The *Fourier transform* \widehat{P} of P at the representation ρ is defined to be the matrix

$$\widehat{P}(\rho) := \sum_{g \in G} P(g) \rho(g).$$

Suppose that P and Q are functions on G . The *convolution* $P * Q$ of P and Q is defined, for all $g \in G$, by

$$P * Q(g) := \sum_{h \in G} P(gh^{-1}) Q(h).$$

The Fourier transform converts the convolution of functions P and Q into the (argument-wise) multiplication of their transforms \widehat{P} and \widehat{Q} , i.e.,

$$\widehat{P * Q}(\rho) = \widehat{P}(\rho)\widehat{Q}(\rho).$$

In the special case when P is a class function, it is a consequence of Schur's Lemma (2.3.3) that the Fourier transform may be calculated easily by use of the following, which is Lemma 5 of Diaconis and Shahshahani (1981).

LEMMA 2.5.1. *Suppose that ρ is an irreducible representation of a finite group G with character χ and that P is a class function. For each conjugacy class i , let P_i be the constant value of P on the class, let n_i be the cardinality of the class, and let χ_i be the constant value of χ on the class. Then the Fourier transform of P is given by*

$$\widehat{P}(\rho) = \left[\frac{1}{d_\rho} \sum_{i=1}^s P_i n_i \chi_i \right] I,$$

where d_ρ is the dimension of ρ , I is the d_ρ -dimensional identity matrix, and the sum is taken over distinct conjugacy classes.

The Fourier transforms of any distribution at the trivial representation and of the uniform distribution at any nontrivial representation have special forms, as described in the following, which is an immediate consequence of the preceding lemma and Proposition 2.4.4.

COROLLARY 2.5.2. *Suppose that P is any probability measure and that U is the uniform probability measure, both defined on a finite group G . Then $\widehat{P}(\rho_0) = 1$ at the trivial representation ρ_0 of G and $\widehat{U}(\rho)$ is the $d_\rho \times d_\rho$ zero matrix for each nontrivial representation ρ of G .*

The function P on G may be reconstructed from its Fourier transform by use of the following, which is the *Fourier inversion formula* in Section 6.2 of Serre (1977).

LEMMA 2.5.3. *Suppose that P is a function on G and that \widehat{P} is its Fourier transform. Then for each $g \in G$*

$$P(g) = \frac{1}{|G|} \sum_{i=1}^s d_{\rho_i} \operatorname{tr} \left(\rho_i(g^{-1}) \widehat{P}(\rho_i) \right),$$

where $\rho_1, \rho_2, \dots, \rho_s$ are the nonisomorphic irreducible representations of G with dimensions $d_{\rho_1}, d_{\rho_2}, \dots, d_{\rho_s}$, respectively.

A consequence of this formula is another useful result, which is the *Plancherel formula* in Section 6.2 of Serre (1977).

LEMMA 2.5.4. *Suppose that P and Q are functions on G and that \widehat{P} and \widehat{Q} are their Fourier transforms. Then*

$$\sum_{g \in G} P(g)Q(g^{-1}) = \frac{1}{|G|} \sum_{i=1}^s d_{\rho_i} \operatorname{tr} \left(\widehat{P}(\rho_i) \widehat{Q}(\rho_i) \right),$$

where $\rho_1, \rho_2, \dots, \rho_s$ are the nonisomorphic irreducible representations of G with dimensions $d_{\rho_1}, d_{\rho_2}, \dots, d_{\rho_s}$, respectively.

2.6. Random Walks on Groups. We now turn our attention to the subject of random walks on groups. A more detailed introduction to this subject may be found in Chapter 3 of Diaconis (1988).

Suppose that P is a probability measure defined on a group G . Let ξ_1, ξ_2, \dots be a sequence of independent G -valued random variables each distributed according to P . A *random walk* on G is a sequence $\mathbf{X} = (X_0, X_1, X_2, \dots)$ defined by $X_0 := e \in G$ and $X_n := \xi_n \xi_{n-1} \cdots \xi_1$ for all $n \geq 1$. Notice that \mathbf{X} is a Markov chain with state space G :

$$\begin{aligned} & \mathbb{P}\{X_n = g_n \mid X_0 = e, X_1 = g_1, \dots, X_{n-1} = g_{n-1}\} \\ &= \mathbb{P}\{\xi_n \xi_{n-1} \cdots \xi_1 = g_n \mid \xi_1 = g_1, \dots, \xi_{n-1} \cdots \xi_1 = g_{n-1}\} \\ &= \mathbb{P}\{\xi_n = g_n g_{n-1}^{-1}\} = P(g_n g_{n-1}^{-1}) \end{aligned}$$

for all $n \geq 1$ and all $g_1, \dots, g_n \in G$. In this way a probability measure P on G induces a transition matrix \mathbf{P} , where the entry of \mathbf{P} at the intersection of the row corresponding to $g \in G$ and the column corresponding to $h \in G$ is given by

$$\mathbf{P}_{gh} = P(hg^{-1}).$$

In the special case where P is a class function, we may determine all the eigenvalues of the transition matrix \mathbf{P} , together with their multiplicities, by use of the following, which is Corollary 3 of Diaconis and Shahshahani (1981).

LEMMA 2.6.1. *Suppose P is a probability measure defined on a finite group G and that P is a class function. Let \mathbf{P} be the transition matrix of the Markov chain induced by the probability measure P . Then, for each irreducible representation ρ of G , there is an eigenvalue π_ρ of \mathbf{P} occurring with algebraic multiplicity d_ρ^2 such that*

$$\pi_\rho = \frac{1}{d_\rho} \sum_{i=1}^s P_i n_i \chi_i,$$

where the sum is taken over distinct conjugacy classes.

Notice that $\pi_\rho \cdot I$ in the lemma above is exactly the value of $\widehat{P}(\rho)$ in Lemma 2.5.1.

In order to discuss the convergence of these random walks to their stationary distributions, we need a metric between probability measures. Suppose that P and Q are two probability measures defined on a finite group G . The *total variation distance* between P and Q is defined by

$$\|P - Q\|_{\text{TV}} := \max_{A \subseteq G} |P(A) - Q(A)|,$$

while the ℓ^1 distance between P and Q is defined by

$$\|P - Q\|_1 := \sum_{g \in G} |P(g) - Q(g)|.$$

Notice that $\|P - Q\|_{\text{TV}} = \frac{1}{2} \|P - Q\|_1$. The ℓ^2 distance between P and Q is defined by

$$\|P - Q\|_2 := \left(\sum_{g \in G} |P(g) - Q(g)|^2 \right)^{1/2}.$$

All three of these measures of distance are indeed metrics. It is a direct consequence of the Cauchy-Schwarz inequality that

$$\|P - Q\|_{\text{TV}}^2 \leq \frac{1}{4} |G| \|P - Q\|_2^2.$$

We are now able to bound the distance to uniformity of a probability measure P in terms of its Fourier transform by use of the following, which is the Upper Bound Lemma in Section B of Chapter 3 of Diaconis (1988).

LEMMA 2.6.2. *Suppose that P is a probability measure defined on a finite group G . Then*

$$\|P - U\|_{\text{TV}}^2 \leq \frac{1}{4} |G| \cdot \|P - U\|_2^2 = \frac{1}{4} \sum_{\rho} d_\rho \operatorname{tr} \left(\widehat{P}(\rho) \widehat{P}(\rho)^* \right)$$

where the sum is taken over all nontrivial irreducible representations of G and $\widehat{P}(\rho)^*$ is the conjugate transpose of $\widehat{P}(\rho)$.

In the Upper Bound Lemma (2.6.2), the inequality follows from the Cauchy-Schwarz inequality; the equality follows from the Plancherel formula (2.5.4) and Corollary 2.5.2.

2.7. Random Walk on the Symmetric Group. We now return to the random walk on the symmetric group defined in Section 2.1 and the proof of Theorem 2.1.3. We include only those portions of the proof that will be needed to conduct our analysis in later sections. A detailed introduction to the symmetric group and its representation theory may be found in Chapters 1 and 2 of James and Kerber (1981). Another source is Sagan (1991).

For any $\pi \in S_n$, there is an associated n -dimensional vector $a = (a_1, a_2, \dots, a_n)$, called the *cycle type* of π , where a_i is the numbers of cycle factors of π of length i for $1 \leq i \leq n$. According to Lemma 1.2.6 of James and Kerber (1981), two elements of S_n are conjugate if and only if their cycle types are identical; hence, there is a one-to-one correspondence between these cycle type vectors and conjugacy classes of S_n . Thus, since the transpositions in S_n form their own conjugacy class, the probability measure P defined in (2.1.1) is a class function.

A vector $[\lambda] = [\lambda_1, \lambda_2, \dots, \lambda_k]$ such that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$ and $\lambda_1 + \lambda_2 + \dots + \lambda_k = n$ is called a *partition* of n . According to Theorem 2.1.11 of James and Kerber (1981), there is a one-to-one correspondence between nonisomorphic irreducible representations of S_n and partitions $[\lambda]$ of n . Thus we have identified all of the irreducible representations of S_n , over which the summation is taken in the Upper Bound Lemma (2.6.2).

Lemmas 2.5.1 and 2.4.1(a) are then used to calculate the Fourier transform of an irreducible representation $\rho_{[\lambda]}$ of S_n , which is

$$\widehat{P}(\lambda) = \left[\frac{1}{n} + \frac{n-1}{n} r(\lambda) \right] I,$$

where $r(\lambda) := \chi_{[\lambda]}(\tau)/d_{[\lambda]}$, $\chi_{[\lambda]}(\tau)$ is the character of $\rho_{[\lambda]}$ at any transposition τ , $d_{[\lambda]}$ is the dimension of $\rho_{[\lambda]}$, and I is the $d_{[\lambda]}$ -dimensional identity matrix.

It then follows from the Upper Bound Lemma (2.6.2) that

$$\|P^{*k} - U\|_{\text{TV}}^2 \leq \frac{1}{4} n! \|P^{*k} - U\|_2^2 = \frac{1}{4} \sum_{[\lambda]} d_{[\lambda]}^2 \left[\frac{1}{n} + \frac{n-1}{n} r(\lambda) \right]^{2k}$$

where the term not to be included in the summation occurs when $[\lambda] = [n]$, which corresponds to the trivial representation of S_n .

The following formulas, found in Section D of Chapter 3 and Section B of Chapter 7, respectively, of Diaconis (1988) are used to calculate the numerical value of the Fourier transform.

LEMMA 2.7.1. *Suppose that ρ is an irreducible representation of S_n corresponding to the partition $[\lambda] = [\lambda_1, \dots, \lambda_k]$ of n . Let $r(\lambda) := \chi_{[\lambda]}(\tau)/d_{[\lambda]}$ with $\tau \in S_n$. Then*

$$r(\lambda) = \frac{1}{n(n-1)} \sum_{j=1}^k [\lambda_j^2 - (2j-1)\lambda_j] \quad \text{and}$$

$$d_{[\lambda]} = n! \det \left(\frac{1}{(\lambda_i - i + j)!} \right)_{1 \leq i, j \leq k},$$

with $1/m! := 0$ if $m < 0$.

A lengthy, detailed discussion in Section D of Chapter 3 of Diaconis (1988) determines the existence of a universal constant $a > 0$ such that if $k = \frac{1}{2}n \log n + cn$ with $c > 0$, then

$$\frac{1}{4} \sum_{[\lambda]} d_{[\lambda]}^2 \left[\frac{1}{n} + \frac{n-1}{n} r(\lambda) \right]^{2k} \leq a^2 e^{-4c}. \quad (2.7.2)$$

This completes the proof of Theorem 2.1.3.

Theorem 2.1.3 shows that $k = \frac{1}{2}n \log n + cn$ steps are sufficient for the (normalized) ℓ^2 distance, and hence the total variation distance, to become small. That $k = \frac{1}{2}n \log n - cn$ steps are also *necessary* is a result of the following, which is a solution to Exercise 13 in Section D of Chapter 3 of Diaconis (1988). The method employed in the proof is a standard technique used in problems of this sort.

THEOREM 2.7.3. *Let P and U be the probability measures on the symmetric group S_n defined in (2.1.1) and (2.1.2), respectively. Let $k = \frac{1}{2}n \log n - cn$ be a nonnegative integer, with $c > 0$. Then there exists a universal constant $\tilde{a} > 0$ such that*

$$\frac{1}{2} (n!)^{1/2} \|P^{*k} - U\|_2 \geq \|P^{*k} - U\|_{\text{TV}} \geq 1 - \tilde{a}e^{-2c}.$$

PROOF. Let χ be the character of the representation $\rho_{[n-1,1]}$ of S_n ; this representation corresponds to the largest term in the summation from the proof of Theorem 2.1.3. Since any element of S_n is conjugate to its inverse, it follows from parts (b) and (c) of Lemma 2.4.1 that χ is real.

Under the uniform measure U , it follows from Proposition 2.4.4 that

$$E_U(\chi) = \frac{1}{n!} \sum_{\pi \in S_n} \chi(\pi) = \langle \chi, \chi_0 \rangle_{S_n} = 0,$$

where χ_0 is the character of the trivial representation $\rho_{[n]}$, and that

$$\text{Var}_U(\chi) = E_U(\chi^2) = \frac{1}{n!} \sum_{\pi \in S_n} \chi(\pi)^2 = \langle \chi, \chi \rangle_{S_n} = 1.$$

It follows from Lemma 2.7.1 that $d_{[n-1,1]} = n - 1$ and $r([n - 1, 1]) = \frac{n-3}{n-1}$. Thus under the k -fold convolution measure P^{*k} , it follows from Lemma 2.5.1 and (the calculations in the proof in Section 2.7 of) Theorem 2.1.3 that

$$E_{P^{*k}}(\chi) = \sum_{\pi \in S_n} P^{*k}(\pi) \chi(\pi) = \text{tr} \sum_{\pi \in S_n} P^{*k}(\pi) \rho(\pi) = \text{tr} \widehat{P^{*k}}(\rho) = (n - 1) \left(1 - \frac{2}{n}\right)^k,$$

where $\rho = \rho_{[n-1,1]}$.

In order to determine $\text{Var}_{P^{*k}}(\chi)$, we must now calculate $E_{P^{*k}}(\chi^2)$. It follows from Lemma 2.4.2 that χ^2 is the character of the representation $\rho_{[n-1,1]} \otimes \rho_{[n-1,1]}$. Recall from Proposition 2.3.4 that every representation is the direct sum of irreducible representations; in this case, it follows from the example following Lemma 2.9.16 of James and Kerber (1981) that we have the isomorphism

$$\rho_{[n-1,1]} \otimes \rho_{[n-1,1]} \cong \rho_{[n]} \oplus \rho_{[n-1,1]} \oplus \rho_{[n-2,2]} \oplus \rho_{[n-2,1,1]}.$$

So it follows also from Lemma 2.4.2 that

$$E_{P^{*k}}(\chi^2) = E_{P^{*k}}(\chi_{[n]}) + E_{P^{*k}}(\chi_{[n-1,1]}) + E_{P^{*k}}(\chi_{[n-2,2]}) + E_{P^{*k}}(\chi_{[n-2,1,1]}).$$

As above, it follows from Lemma 2.7.1 that $d_{[n-2,2]} = \frac{1}{2}n(n-3)$ and $r([n-2,2]) = \frac{n-4}{n}$ and that $d_{[n-2,1,1]} = \frac{1}{2}(n-1)(n-2)$ and $r([n-2,1,1]) = \frac{n-5}{n-1}$. Thus under the k -fold convolution measure P^{*k} , it follows from Lemma 2.5.1 and (the calculations in the proof in Section 2.7 of) Theorem 2.1.3 that

$$E_{P^{*k}}(\chi_{[n]}) = 1,$$

$$E_{P^{*k}}(\chi_{[n-2,2]}) = \frac{1}{2}n(n-3) \left(1 - \frac{2}{n}\right)^{2k}, \text{ and}$$

$$E_{P^{*k}}(\chi_{[n-2,1,1]}) = \frac{1}{2}(n-1)(n-2) \left(1 - \frac{4}{n}\right)^k.$$

These results combine to show that

$$\begin{aligned} \text{Var}_{P^{*k}}(\chi) &= 1 + (n-1) \left(1 - \frac{2}{n}\right)^k + \frac{1}{2}(n-1)(n-2) \left(1 - \frac{4}{n}\right)^k \\ &\quad - \frac{1}{2}(n^2 - n + 2) \left(1 - \frac{2}{n}\right)^{2k}. \end{aligned}$$

Let $k = \frac{1}{2}n \log n - cn$. By elementary calculus, $x \leq -\log(1-x) \leq \frac{x}{1-x}$ for $0 \leq x < 1$. Thus, if $n \geq 3$ and $c \geq 0$,

$$\begin{aligned} E_{P^{*k}}(\chi) &= (n-1) \left(1 - \frac{2}{n}\right)^k \geq (n-1) e^{-2k/(n-2)} \\ &= \left(1 - \frac{1}{n}\right) \left(\frac{1}{n}\right)^{2/(n-2)} e^{2c} e^{4c/(n-2)} \geq \frac{2}{27} e^{2c}. \end{aligned}$$

where we note that, for $n \geq 3$, $\left(1 - \frac{1}{n}\right) \left(\frac{1}{n}\right)^{2/(n-2)}$ is increasing and $e^{4c/(n-2)} \geq 1$.

In order to bound the variance, notice that

$$\begin{aligned} \frac{1}{2}(n-1)(n-2) \left(1 - \frac{4}{n}\right)^k &\leq \frac{1}{2}(n-1)(n-2)e^{-4k/n} = \frac{1}{2} \left(1 - \frac{3}{n} + \frac{2}{n^2}\right) e^{4c} \quad \text{and} \\ \frac{1}{2}(n^2 - n + 2) \left(1 - \frac{2}{n}\right)^{2k} &\geq \frac{1}{2}(n^2 - n + 2)e^{-4k/(n-2)} \\ &= \frac{1}{2} \left(1 - \frac{1}{n} + \frac{2}{n^2}\right) \left(\frac{1}{n}\right)^{4/(n-2)} e^{4c} e^{8c/(n-2)}. \end{aligned}$$

Thus

$$\frac{1}{2}(n-1)(n-2) \left(1 - \frac{4}{n}\right)^k - \frac{1}{2}(n^2 - n + 2) \left(1 - \frac{2}{n}\right)^{2k} \leq 0$$

when

$$\frac{1}{2} \left(1 - \frac{3}{n} + \frac{2}{n^2}\right) e^{4c} - \frac{1}{2} \left(1 - \frac{1}{n} + \frac{2}{n^2}\right) \left(\frac{1}{n}\right)^{4/(n-2)} e^{4c} e^{8c/(n-2)} \leq 0.$$

This restriction is equivalent to

$$c \leq \frac{1}{2} \log n - \left(\frac{n-2}{8}\right) \log \left(1 - \frac{3}{n} + \frac{2}{n^2}\right) + \left(\frac{n-2}{8}\right) \log \left(1 - \frac{1}{n} + \frac{2}{n^2}\right).$$

Thus, when $c \in (0, \frac{1}{2} \log n]$,

$$\text{Var}_{P^{*k}}(\chi) \leq 1 + (n-1) \left(1 - \frac{2}{n}\right)^k \leq 1 + (n-1)e^{-2k/n} = 1 + \left(1 - \frac{1}{n}\right) e^{2c} \leq 1 + e^{2c}.$$

Now define $A_\alpha := \{\pi \in S_n : |\chi(\pi)| \leq \alpha\}$. It follows from Chebyshev's inequality that $U(A_\alpha) \geq 1 - \frac{1}{\alpha^2}$ and that $P^{*k}(A_\alpha) \leq \frac{1 + e^{2c}}{(\frac{2}{27}e^{2c} - \alpha)^2}$, provided $0 \leq \alpha < \frac{2}{27}e^{2c}$. Then

$$\frac{1}{2}(n!)^{1/2} \|P^{*k} - U\|_2 \geq \|P^{*k} - U\|_{\text{TV}} \geq 1 - \frac{1}{\alpha^2} - \frac{1 + e^{2c}}{(\frac{2}{27}e^{2c} - \alpha)^2}.$$

Choosing $\alpha = \frac{1}{27}e^{2c}$ shows that

$$\frac{1}{2}(n!)^{1/2} \|P^{*k} - U\|_2 \geq \|P^{*k} - U\|_{\text{TV}} \geq 1 - 729e^{-2c} - 1458e^{-4c} \geq 1 - 2187e^{-2c},$$

which completes the proof. \square

The upper bound in Theorem 2.1.3, taken together with the lower bound in Theorem 2.7.3, gives an example of the so-called ‘‘cutoff phenomenon.’’ The total variation distance after k steps is nearly 1 until k reaches about $\frac{1}{2}n \log n$ and then drops precipitously toward 0, the dropoff occurring on the relatively small scale of n . For further discussion of the cutoff phenomenon see Diaconis (1988).

3. Random Walk on the Complete Monomial Groups.

3.1. Introduction. We now extend the idea of random transpositions of n cards, introduced in Section 2.1, to a set of n decks of m cards each and beyond. Imagine n decks of cards, labeled 1 through n , in sequential order, each with its m cards in sequential order. Independently choose two integers p and q , uniformly from $\{1, 2, \dots, n\}$.

If $p \neq q$, transpose the decks in positions p and q . Then, independently of the choice of p and q and uniformly (i.e., with probability $\frac{1}{|G|} = \frac{1}{m!}$ each), permute the deck terminating in position p by a permutation in $G = S_m$; and independently, also uniformly, permute the deck terminating in position q . This procedure is denoted by $(\vec{v}; \tau)$, where $\tau \in S_n$ is the transposition $(p\ q)$ and the only possible non-identity entries of $\vec{v} \in G^n = S_m^n$ are in positions p and q . The element of \vec{v} in position p (resp., q) is $\pi \in G = S_m$ if the deck terminating in position p (resp., q) is permuted by $\pi \in G = S_m$.

If $p = q$ (which occurs with probability $1/n$), leave the decks in their current positions. Then, again independently and uniformly, permute the deck in position $p = q$ by a permutation in $G = S_m$. If the order of the deck is changed, this action is denoted by $(\vec{u}; e)$, where $e \in S_n$ is the identity permutation and the only non-identity entry of $\vec{u} \in G^n = S_m^n$ is in position $p = q$. If the order of the deck is not changed, then this action is of course the identity, which is denoted by $(\vec{e}; e)$.

If this process is repeated many times, the decks will appear to be in random order, and each deck will appear to be randomly permuted.

The specific example above was for motivational purposes only. In this section we will actually examine a random walk on $G \wr S_n$ for any group G , not just the symmetric group S_m . In this more general setting, the example above is equivalent to beginning with a vector $(e, \dots, e) \in G^n$, where each $e \in G$ is the identity element. Two elements of this vector are then transposed as the decks of cards were above. The transposed elements of this vector are then multiplied by elements of G as the individual decks were permuted above.

We refer to the process on $G \wr S_n$ described above as the *independent shuffles* random walk, retaining use of the word “shuffles” even when G is not necessarily S_m . A similar process, known as the *paired shuffles* random walk, will be introduced in Section 3.7.

In the special case of the generalized symmetric group $\mathbb{Z}_m \wr S_n$, imagine n wheels, each of which may stop at one of m values. The process described above randomly transposes two of the n wheels and then independently spins the transposed wheels. We thus refer to the process in this case as the *independent spins* random walk.

In the special case of the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, imagine n cards, each with an orientation (up or down). The process described above randomly transposes two of the n cards and then independently flips the transposed cards. We thus refer to the process in this case as the *independent flips* random walk. Schoolfield (1998) used the comparison technique to analyze two random walks by comparing them to the independent flips random walk.

The independent shuffles random walk may be modeled formally by a probability measure P on the complete monomial group $G \wr S_n$. Since there are $\binom{n}{2}$ transpositions of n elements, then there are $|G|^2 \cdot \binom{n}{2} = \frac{1}{2}|G|^2 n(n-1)$ elements of the form $(\vec{v}; \tau)$. There are also $n(|G|-1)$ elements of the form $(\vec{u}; e)$ with $\vec{u} \neq \vec{e}$. We may thus define the following probability measure on the set of all elements of $G \wr S_n$:

$$\begin{aligned} P(\vec{e}; e) &= \frac{1}{|G|n}, \\ P(\vec{u}; e) &= \frac{1}{|G|n^2} \quad \text{where } \vec{u} \neq \vec{e} \in G^n, \\ P(\vec{v}; \tau) &= \frac{2}{|G|^2 n^2} \quad \text{where } \vec{v} \in G^n, \\ P(\vec{x}; \pi) &= 0 \quad \text{otherwise,} \end{aligned} \tag{3.1.1}$$

where there is only one non-identity entry of $\vec{u} \in G^n$, and where if $\tau \in S_n$ is the transposition $(p \ q)$ then the only possible non-identity entries of $\vec{v} \in G^n$ are in positions p and q . In the special case of the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, we refer to the elements $(\vec{u}; e)$ as *signed identities* and the elements $(\vec{v}; \tau)$ as *signed transpositions*.

Since there are $|G|^n \cdot n!$ elements in $G \wr S_n$, the uniform probability measure is given by

$$U(\vec{x}; \pi) = \frac{1}{|G|^n \cdot n!} \quad \text{for } (\vec{x}; \pi) \in G \wr S_n. \tag{3.1.2}$$

The following result establishes an upper bound on both the total variation distance and the ℓ^2 distance between P^{*k} and U . We establish an analogous result for the paired shuffles random walk as Theorem 3.7.3.

THEOREM 3.1.3. *Let P and U be the probability measures on the complete monomial group $G \wr S_n$ defined in (3.1.1) and (3.1.2), respectively. Let $k = \frac{1}{2}n \log n + \frac{1}{4}n \log(|G|-1) + cn$. Then there exists a universal constant $b > 0$ such that*

$$\|P^{*k} - U\|_{\text{TV}} \leq \frac{1}{2} (|G|^n n!)^{1/2} \|P^{*k} - U\|_2 \leq b e^{-2c} \quad \text{for all } c > 0.$$

In the following sections we establish the results necessary to prove this theorem and an analogous theorem for the paired shuffles random walk. In Section 3.2 we study the basic

properties of the complete monomial groups. In Section 3.3 we show that the probability measure defined above is constant on conjugacy classes. In Section 3.4 we identify all of the irreducible representations of the complete monomial groups and in Section 3.5 we calculate the characters of these irreducible representations. In Section 3.6 we calculate the Fourier transform of the probability measure defined in (3.1.1) and prove Theorem 3.1.3 along with a matching ℓ^2 lower bound. In Section 3.7 we perform a similar analysis of the paired shuffles random walk; however, the results are quite different in the case that G is nonabelian.

3.2. The Complete Monomial Groups. The *complete monomial group* $G \wr S_n$ is the wreath product of the group G with the symmetric group S_n . Special cases include the *hyperoctahedral group* $\mathbb{Z}_2 \wr S_n$, the *generalized symmetric group* $\mathbb{Z}_m \wr S_n$, and $S_m \wr S_n$. The elements of $G \wr S_n$ may be represented as $(\vec{x}; \pi) \in G^n \times S_n$. It then follows that the order of the complete monomial group is $|G \wr S_n| = |G|^n \cdot n!$.

Each element $(\vec{x}; \pi) \in G \wr S_n$ acts on a vector $\vec{w} \in G^n$ by first permuting its elements according to the permutation π and then left-multiplying the elements of the permuted vector by the elements of \vec{x} , entry by entry; i.e., for any $(\vec{x}; \pi) \in G \wr S_n$ and any $\vec{w} \in G^n$,

$$(\vec{x}; \pi)(\vec{w}) = (x_1 \cdot w_{\pi^{-1}(1)}, \dots, x_n \cdot w_{\pi^{-1}(n)}) \in G^n.$$

In keeping with the decks of cards analogy, each element $(\vec{x}; \pi) \in S_m \wr S_n$ first permutes the n decks of cards according to π and then permutes each deck of cards according to the entry of \vec{x} at its new index.

Considering the actions, described above, of the elements of the complete monomial group on vectors in G^n , it follows that the product of two elements $(\vec{x}; \pi), (\vec{y}; \sigma) \in G \wr S_n$ is given by

$$(\vec{y}; \sigma) \cdot (\vec{x}; \pi) = (y_1 \cdot x_{\sigma^{-1}(1)}, \dots, y_n \cdot x_{\sigma^{-1}(n)}; \sigma\pi).$$

Thus the identity element is $(e, \dots, e; e)$ where $(e, \dots, e) \in G^n$ and e is the identity permutation in S_n . It also follows that the inverse of any element $(\vec{x}; \pi) \in G \wr S_n$ is

$$(\vec{x}; \pi)^{-1} = (x_{\pi(1)}^{-1}, \dots, x_{\pi(n)}^{-1}; \pi^{-1}).$$

There are two very important subgroups of $G \wr S_n$ which should be considered. The subgroup consisting of elements of the form $\{(\vec{x}; e) : \vec{x} \in G^n\}$ is isomorphic to G^n and is a normal subgroup of $G \wr S_n$. The subgroup consisting of elements of the form $\{(e, \dots, e; \pi) : \pi \in S_n\}$ is isomorphic to S_n . Notice that any element of $G \wr S_n$ can be written uniquely as the product of a single element from each of these two subgroups.

3.3. Conjugacy Classes. Recall that any permutation $\pi \in S_n$ can be written as the product of at most n disjoint cyclic factors. (For any index not moved by π , there is considered to be a cyclic factor of length one.) For any $\pi \in S_n$ there is an associated n -dimensional vector, called the *cycle type* of π , which lists the number of disjoint cyclic factors of π of each possible length.

For any $(\vec{x}; \pi) \in G \wr S_n$, the permutation π can be written as the product of disjoint cyclic factors, as described above. Suppose that π is the product of j disjoint cyclic factors of lengths $\{k_1, \dots, k_j\}$, respectively, with $k_1 + \dots + k_j = n$. Then π can be written as

$$\pi = (i_1^{(1)} i_2^{(1)} \dots i_{k_1}^{(1)}) (i_1^{(2)} i_2^{(2)} \dots i_{k_2}^{(2)}) \dots (i_1^{(j)} i_2^{(j)} \dots i_{k_j}^{(j)})$$

where each $i_m^{(\ell)} \in \{1, 2, \dots, n\}$. For each disjoint cyclic factor of π , we may define

$$g_\ell(\vec{x}; \pi) = x_{i_1^{(\ell)}} \cdot x_{i_2^{(\ell)}} \cdot \dots \cdot x_{i_{k_\ell}^{(\ell)}},$$

which is called the ℓ th *cycle product* of $(\vec{x}; \pi)$. Notice that $g_\ell(\vec{x}; \pi) \in G$ for all $1 \leq \ell \leq j$.

Suppose that G has s conjugacy classes C_1, C_2, \dots, C_s , with C_1 being the conjugacy class of the identity element $e \in G$. By calculating $g_\ell(\vec{x}; \pi)$ for all $1 \leq \ell \leq j$, we may construct an $s \times n$ *type matrix* $a(\vec{x}; \pi)$ for $(\vec{x}; \pi)$ in the following manner. Let the k th entry of the j th row be the number of cyclic factors of length k contained in π for which $g_\ell(\vec{x}; \pi) \in C_j$. According to Theorem 4.2.8 of James and Kerber (1981), two elements of $G \wr S_n$ are conjugate if and only if their type matrices are identical. Notice that the vector of column sums of $a(\vec{x}; \pi)$ is the cycle type vector of π .

We are primarily interested in the type matrices of the elements in the support of the probability measure defined in (3.1.1). In the identity element $(\vec{e}; e) \in G \wr S_n$, the identity permutation $e \in S_n$ is the product of n disjoint cyclic factors each of length one, and each of these cyclic factors corresponds to an identity element in $\vec{e} = (e, \dots, e) \in G^n$. Thus the type matrix of the identity element is the $s \times n$ matrix with $a_{11}(\vec{e}; e) = n$ and the remaining entries all zeros.

In the special case of the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, we have $s = 2$. Thus the type matrix of the identity element is

$$a(\vec{0}; e) = \begin{pmatrix} n & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Notice that in each $(\vec{u}; e) \in G \wr S_n$ with $\vec{u} \neq \vec{e}$, the identity permutation $e \in S_n$ is again the product of n disjoint cyclic factors each of length one. However, $\vec{u} \in G^n$ has a single non-identity entry, which corresponds to one of the cyclic factors of the identity permutation. Thus, there are $s - 1$ different type matrices for these elements, each with $a_{11}(\vec{u}; e) = n - 1$,

with $a_{k1}(\vec{u}; e) = 1$ if the non-identity entry of \vec{u} is in C_k and the remaining entries all zeros. Hence, the elements of the form $(\vec{u}; e) \in G \wr S_n$ split into $s - 1$ conjugacy classes.

In the special case of the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, the type matrix of each of the signed identities, which together form a single conjugacy class, is

$$a(\vec{u}; e) = \begin{pmatrix} n-1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

In each $(\vec{v}; \tau) \in G \wr S_n$, the transposition $\tau \in S_n$ is the product of $n - 2$ cyclic factors of length one and one cyclic factor of length two. Also, each of the $n - 2$ cyclic factors of length one corresponds to an identity $e \in G$ in $\vec{v} \in G^n$, while the only possible non-identity entries in $\vec{v} \in G^n$ correspond to the one cyclic factor of length two. Thus $g_\ell(\vec{v}; \tau) = 0$ for each of the $n - 2$ cyclic factors of length one, but $g_\ell(\vec{v}; \tau)$ can be an arbitrary element of G for the one cyclic factor of length two. Thus, there are s different type matrices for these elements, each with $a_{11}(\vec{v}; \tau) = n - 2$, with $a_{k2}(\vec{v}; \tau) = 1$ if $g_\ell(\vec{v}; \tau) \in C_k$ for the one cyclic factor of length two, and with the remaining entries all zeros. Hence, the elements of the form $(\vec{v}; \tau) \in G \wr S_n$ split into s conjugacy classes.

In the special case of the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, the calculation of the type matrix splits the signed transpositions into two sets. Let us refer to a signed transposition which flips neither or both of the cards as an *even* transposition and designate it as $(\vec{v}; \tau^+)$. We will refer to one which flips either, but not both, of the cards as an *odd* transposition and designate it as $(\vec{v}; \tau^-)$. Thus the type matrices of the even transpositions and the odd transpositions are, respectively,

$$a(\vec{v}; \tau^+) = \begin{pmatrix} n-2 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad a(\vec{v}; \tau^-) = \begin{pmatrix} n-2 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{pmatrix}.$$

We have now identified the type matrices (and, hence, the conjugacy classes) of all the elements in the support of the probability measure defined in (3.1.1). Furthermore, we have established the following.

LEMMA 3.3.1. *The probability measure defined in (3.1.1) is constant on conjugacy classes.*

Knowing that there are the same number of conjugacy classes as there are type matrices allows us to calculate the number of conjugacy classes. According to Lemma 4.2.9 of James and Kerber (1981), the number of conjugacy classes of $G \wr S_n$ is equal to

$$\sum_{(n_1, n_2, \dots, n_s)} p(n_1) \cdot p(n_2) \cdots p(n_s)$$

where $p(k)$ is the number of partitions of k (with $p(0) := 1$) and the sum is taken over all s -dimensional vectors such that $n_1 + n_2 + \cdots + n_s = n$ and $n_j \geq 0$ for all $1 \leq j \leq s$.

We may also calculate the order of each of the conjugacy classes. According to Lemma 4.2.10 of James and Kerber (1981), the number of elements of a particular type $a(\vec{x}; \pi)$ in $G \wr S_n$, and hence the number of elements of a particular conjugacy class, is equal to

$$\frac{|G|^n \cdot n!}{\prod_{i,j} [(j|G|/|C_i|)^{a_{ij}} \cdot a_{ij}!]}$$

where a_{ij} is the (i, j) th element of $a(\vec{x}; \pi)$ and $|C_i|$ is the order of the i th conjugacy class of G . By applying this result to the elements of $G \wr S_n$ whose type matrices were determined above, we have the following otherwise obvious corollary.

COROLLARY 3.3.2. *There are $n|C_k|$ elements of the form $(\vec{u}; e)$, for $2 \leq k \leq s$, and there are $\frac{1}{2}n(n-1) \cdot |G| \cdot |C_k|$ elements of the form $(\vec{v}; \tau)$, for $1 \leq k \leq s$.*

3.4. Irreducible Representations. We now construct a collection of irreducible representations of the complete monomial groups from irreducible representations of certain of their subgroups. We will later see that *every* irreducible representation of $G \wr S_n$ can be constructed in such a manner. The method used is from Section 4.3 of James and Kerber (1981). Other methods described in Section 8.2 of Serre (1977) and in Chapter V of Simon (1996) could be used in the special case when G is abelian.

Let G^* be the subgroup of $G \wr S_n$ consisting of all elements of the form $\{(\vec{x}; e) : \vec{x} \in G^n\}$ and let H^* be the subgroup of $G \wr S_n$ consisting of all elements of the form $\{(e, \dots, e; \pi) : \pi \in S_n\}$. Recall from Section 3.2 that G^* is a normal subgroup of $G \wr S_n$ and that $G^* \cong G^n$ and $H^* \cong S_n$. Recall, furthermore, that each element of $G \wr S_n$ can be written uniquely as a product gh with $g \in G^*$ and $h \in H^*$.

Suppose that G has s conjugacy classes. Then it follows from Proposition 2.3.1 that there are s irreducible representations of G , namely, $\rho_1, \rho_2, \dots, \rho_s$, where we choose our labelling so that ρ_1 is the trivial representation. So it follows from Proposition 2.3.5 that any irreducible representation of G^n is isomorphic to an n -fold tensor product

$$\rho^{(1)} \otimes \rho^{(2)} \otimes \cdots \otimes \rho^{(n)}$$

where each $\rho^{(i)} \in \{\rho_1, \rho_2, \dots, \rho_s\}$. Notice that any irreducible representation of G^n can be represented by a vector $\vec{x} \in \{1, \dots, s\}^n$, where the entries of \vec{x} correspond to the indices of the factors in the tensor product above.

Suppose that the tensor product forming a particular irreducible representation of G^n contains n_j factors of the kind ρ_j , for all $1 \leq j \leq s$. The vector $(n) = (n_1, n_2, \dots, n_s)$ is called the *type* of this representation. Notice that for each type $(n) = (n_1, n_2, \dots, n_s)$, there is a unique representative whose first n_1 factors in the tensor product are ρ_1 , whose next n_2 factors are ρ_2 , etc. We will refer to this canonical representative as $\rho_{(n)}$.

Since $G^n \cong G^* = \{(\vec{x}; e) : \vec{x} \in G^n\}$, any representation ρ of G^n extends easily to a representation of G^* by setting $\rho(\vec{x}; e) \equiv \rho(\vec{x})$. Thus any representation of G^n is also a representation of G^* . Therefore, the collection $\{\rho_{(n)}\}$ is a complete system of irreducible representations of G^* which are of different types.

We now turn our attention to $H^* = \{(e, \dots, e; \pi) : \pi \in S_n\} \cong S_n$. For each type $(n) = (n_1, n_2, \dots, n_s)$, define $S_{(n)}$ to be the subgroup of S_n which permutes the first n_1 indices among themselves, the next n_2 indices among themselves, etc.; but does not commingle these s sets of indices. Thus $S_{(n)} \cong S_{n_1} \times S_{n_2} \times \dots \times S_{n_s}$, where we define $S_{n_j} := S_1$ if $n_j = 0$ for any $1 \leq j \leq s$.

Recall that there is a one-to-one correspondence between irreducible representations of S_n and partitions $[\lambda]$ of n , where $[\lambda] = [\lambda_1, \lambda_2, \dots, \lambda_k]$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$ and $\lambda_1 + \lambda_2 + \dots + \lambda_k = n$. Thus any irreducible representation of S_n may be denoted as $\rho_{[\lambda]}$, where $[\lambda]$ is the corresponding partition of n .

It follows from Proposition 2.3.5 that any irreducible representation of $S_{n_1} \times S_{n_2} \times \dots \times S_{n_s}$ is isomorphic to the s -fold tensor product

$$\rho_{[\lambda_1]} \otimes \rho_{[\lambda_2]} \otimes \dots \otimes \rho_{[\lambda_s]}$$

where $(\lambda) = ([\lambda_1], [\lambda_2], \dots, [\lambda_s])$ are partitions of $\{n_1, n_2, \dots, n_s\}$ and $\{\rho_{[\lambda_1]}, \rho_{[\lambda_2]}, \dots, \rho_{[\lambda_s]}\}$ are irreducible representations of $\{S_{n_1}, S_{n_2}, \dots, S_{n_s}\}$, respectively. We will denote this tensor product as $\rho_{(\lambda)}$. Thus, there is a one-to-one correspondence between irreducible representations of $S_{n_1} \times S_{n_2} \times \dots \times S_{n_s}$ and ordered s -tuples of partitions of $(n) = (n_1, n_2, \dots, n_s)$. Since $S_{n_1} \times S_{n_2} \times \dots \times S_{n_s} \cong S_{(n)}$, then any representation of $S_{n_1} \times S_{n_2} \times \dots \times S_{n_s}$ is also a representation of $S_{(n)}$.

Let us define $S_{(n)}^* = \{(e, \dots, e; \pi) : \pi \in S_{(n)}\}$. Since $S_{(n)}$ is a subgroup of S_n , it follows that $S_{(n)}^*$ is a subgroup of $G^n \wr S_n$. Since $S_{(n)}^* \cong S_{(n)}$, any representation ρ of $S_{(n)}$ extends easily to a representation of $S_{(n)}^*$ by setting $\rho(e, \dots, e; \pi) \equiv \rho(\pi)$. Thus any representation of $S_{(n)}$ is also a representation of $S_{(n)}^*$. Therefore, for fixed type $(n) = (n_1, n_2, \dots, n_s)$, the collection $\{\rho_{(\lambda)}\}$ is a complete system of irreducible representations of $S_{(n)}^*$, when (λ) ranges over all ordered s -tuples of partitions of (n_1, n_2, \dots, n_s) , respectively.

Now consider the wreath product $G \wr S_{(n)}$, which is a subgroup of $G \wr S_n$. The representations $\rho_{(n)} : G^* \longrightarrow \mathbf{GL}(V_1)$ and $\rho_{(\lambda)} : S_{(n)}^* \longrightarrow \mathbf{GL}(V_2)$, where $(\lambda) = ([\lambda_1], [\lambda_2], \dots, [\lambda_s])$ are partitions of $(n) = (n_1, n_2, \dots, n_s)$, respectively, may be combined to form representations $\rho_{(n)} \otimes \rho_{(\lambda)}$ of $G \wr S_{(n)}$ by defining (with an abuse of the notation “ \otimes ,” since this is not quite the same notion of tensor product as in Section 2.3)

$$(\rho_{(n)} \otimes \rho_{(\lambda)}) (\vec{x}; \pi)(v_1 \otimes v_2) := \rho_{(n)}(\vec{x}; e)(v_1) \otimes \rho_{(\lambda)}(\vec{0}; \pi)(v_2)$$

for all $(\vec{x}; \pi) \in G \wr S_{(n)}$ and all $v_1 \in V_1$ and $v_2 \in V_2$. These representations $\{\rho_{(n)} \otimes \rho_{(\lambda)}\}$ of $G \wr S_{(n)}$ may now be used to induce representations of $G \wr S_n$. More importantly, we have the following, which is Theorem 4.4.3 of James and Kerber (1981).

LEMMA 3.4.1. *The collection of representations of $G \wr S_n$ induced by the collection $\{\rho_{(n)} \otimes \rho_{(\lambda)}\}$ of representations of $G \wr S_{(n)}$ is a complete collection of pairwise inequivalent and irreducible representations of $G \wr S_n$ if $(n) = (n_1, n_2, \dots, n_s)$ ranges over all different types and, for fixed type (n) , $(\lambda) = ([\lambda_1], [\lambda_2], \dots, [\lambda_s])$ ranges over all ordered s -tuples of partitions of (n_1, n_2, \dots, n_s) , respectively.*

Since the number of irreducible representations equals the number of conjugacy classes (Proposition 2.3.1), it follows from the results in Section 3.3 that the number of irreducible representations of $G \wr S_n$ is equal to

$$\sum_{(n_1, n_2, \dots, n_s)} p(n_1) \cdot p(n_2) \cdots p(n_s),$$

where $p(k)$ is the number of partitions of k (with $p(0) := 1$) and the sum is taken over all s -dimensional vectors such that $n_1 + n_2 + \cdots + n_s = n$ and $n_j \geq 0$ for all $1 \leq j \leq s$. This is Corollary 4.4.4 of James and Kerber (1981) and is consistent with the results found above.

3.5. Irreducible Characters. For the elements in the support of the probability measure defined in (3.1.1), we now determine the characters of the irreducible representations of the complete monomial group $G \wr S_n$ induced by the irreducible representations of $G \wr S_{(n)}$ found in Lemma 3.4.1. We do this with the aid of Lemma 2.4.3. The Frobenius character formula, which is Theorem V.4.1 of Simon (1996) or Theorem 12 of Serre (1977), may also be used.

For notational purposes, let $C_k^{(\vec{u}; e)}$ (with $2 \leq k \leq s$) be the conjugacy classes of $(\vec{u}; e)$ in $G \wr S_n$, with k chosen so that the single non-identity entry of \vec{u} is in conjugacy class C_k of G , and let $C_k^{(\vec{v}; \tau)}$ (with $1 \leq k \leq s$) be the conjugacy classes of $(\vec{v}; \tau)$ in $G \wr S_n$, with k chosen

so that the product (calculated as in the determination of the type matrices in Section 3.3) of the two possible non-identity entries of \vec{v} is in conjugacy class C_k of G .

LEMMA 3.5.1. *For the elements in the support of the probability measure defined in (3.1.1), the character of the irreducible representation ρ of the complete monomial group $G \wr S_n$ induced by the irreducible representation $\rho_{(n)} \otimes \rho_{(\lambda)}$ of $G \wr S_{(n)}$ is given by*

$$\begin{aligned}\chi_\rho(\vec{e}; e) &= \binom{n}{n_1, \dots, n_s} d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s} \cdot d_{[\lambda_1]} \cdots d_{[\lambda_s]} = d_\rho, \\ \chi_\rho(\vec{u}; e) &= d_\rho \sum_{j=1}^s \binom{n_j}{n} \left[\frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}} \right] \quad \text{for } (\vec{u}; e) \in C_k^{(\vec{u}; e)}, \\ \chi_\rho(\vec{v}; \tau) &= d_\rho \sum_{j=1}^s \left[\frac{n_j(n_j - 1)}{n(n - 1)} \right] \cdot \left[\frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}} \right] \cdot r(\lambda_j) \quad \text{for } (\vec{v}; \tau) \in C_k^{(\vec{v}; \tau)},\end{aligned}$$

where, for $1 \leq j \leq s$, χ_{ρ_j} is the character of the irreducible representation ρ_j of G and $\chi_{[\lambda_j]}$ is the character of the irreducible representation $\rho_{[\lambda_j]}$ of S_{n_j} , where $r(\lambda_j) := \chi_{[\lambda_j]}(\tau)/d_{[\lambda_j]}$ with transposition $\tau \in S_{n_j}$, and where g_k is any element of the conjugacy class C_k of G .

PROOF. Let $\tilde{G} := G \wr S_n$ and $\tilde{H} := G \wr S_{(n)}$; we apply Lemma 2.4.3 to these groups. Notice that $|\tilde{G}| = |G \wr S_n| = |G|^n \cdot n!$ and that $|\tilde{H}| = |G \wr S_{(n)}| = |G \wr (S_{n_1} \times \cdots \times S_{n_s})| = |G|^n \cdot n_1! \cdots n_s!$. Thus

$$\frac{|\tilde{G}|}{|\tilde{H}|} = \frac{|G|^n \cdot n!}{|G|^n \cdot n_1! \cdots n_s!} = \binom{n}{n_1, \dots, n_s}.$$

Recall, in the representation $\rho_{(n)} \otimes \rho_{(\lambda)}$ of $\tilde{H} = G \wr S_{(n)}$, that $\rho_{(n)}$ is the n -fold tensor product

$$\rho^{(1)} \otimes \rho^{(2)} \otimes \cdots \otimes \rho^{(n)},$$

where each $\rho^{(i)} \in \{\rho_1, \rho_2, \dots, \rho_s\}$ is an irreducible representation of G and the first n_1 factors are ρ_1 , the next n_2 factors are ρ_2 , etc. Also recall that $\rho_{(\lambda)}$ is

$$\rho_{[\lambda_1]} \otimes \rho_{[\lambda_2]} \otimes \cdots \otimes \rho_{[\lambda_s]},$$

where $\rho_{[\lambda_1]}, \rho_{[\lambda_2]}, \dots, \rho_{[\lambda_s]}$ are irreducible representations of $S_{n_1}, S_{n_2}, \dots, S_{n_s}$, respectively. It then follows from Lemma 2.4.2 that the character of the representation $\rho_{(n)} \otimes \rho_{(\lambda)}$ is given by

$$\chi_{\tilde{H}} = \chi^{(1)} \cdots \chi^{(n)} \cdot \chi_{[\lambda_1]} \cdots \chi_{[\lambda_s]}.$$

We begin with the identity element $(\vec{e}; e) \in \tilde{G} = G \wr S_n$. Since the identity element forms a singleton conjugacy class, we have $\ell = 1$. Furthermore, since the only element of \tilde{H} to which it is conjugate in \tilde{G} is the identity element $(\vec{e}; e, \dots, e) \in \tilde{H} = G \wr S_{(n)}$, we have $t = 1$ and $k_1 = 1$.

Recall that the character at the identity of any representation is the dimension of the representation. Thus

$$\chi^{(1)}(e) \cdots \chi^{(n)}(e) = d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s}.$$

So the character of the identity element $(\vec{e}; e, \dots, e) \in \tilde{H} = G \wr S_{(n)}$ is

$$d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s} \cdot \chi_{[\lambda_1]}(e) \cdots \chi_{[\lambda_s]}(e).$$

Therefore, it follows from Lemma 2.4.3 that the character at the identity element in the induced irreducible representation ρ of $G \wr S_n$ is

$$\chi_\rho(\vec{e}; e) = \binom{n}{n_1, \dots, n_s} d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s} \cdot \chi_{[\lambda_1]}(e) \cdots \chi_{[\lambda_s]}(e),$$

and, furthermore, that the dimension of the induced irreducible representation ρ of $G \wr S_n$ is

$$d_\rho = \binom{n}{n_1, \dots, n_s} d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s} \cdot d_{[\lambda_1]} \cdots d_{[\lambda_s]}.$$

We now consider the elements $(\vec{u}; e) \in \tilde{G} = G \wr S_n$ with $\vec{u} \neq \vec{e}$, which comprise $s - 1$ conjugacy classes: $C_2^{(\vec{u}; e)}, C_3^{(\vec{u}; e)}, \dots, C_s^{(\vec{u}; e)}$, where the indices are chosen so that, for $C_k^{(\vec{u}; e)}$, the single non-identity entry of \vec{u} is in conjugacy class C_k of G . For a particular conjugacy class $C_k^{(\vec{u}; e)}$, since there are $n|C_k|$ elements in \tilde{G} , we have $\ell = n|C_k|$. Each of the $s - 1$ conjugacy classes splits into s classes in \tilde{H} : in the first class, the only non-identity entry of \vec{u} is one of its first n_1 entries; in the second class, the only non-identity entry of \vec{u} is one of its next n_2 entries; etc. Thus $t = s$ with $k_j = n_j|C_k|$ for all $1 \leq j \leq s$.

For a particular conjugacy class C_k of G , let g_k be a representative element. Then the characters of the s classes of elements $(\vec{u}; e, \dots, e) \in \tilde{H} = G \wr S_{(n)}$ are

$$d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s} \cdot d_{[\lambda_1]} \cdots d_{[\lambda_s]} \cdot \frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}},$$

for $1 \leq j \leq s$.

Therefore, it follows from Lemma 2.4.3 that the character at $(\vec{u}; e)$ of the induced irreducible representation ρ of $G \wr S_n$ is

$$\chi_\rho(\vec{u}, e) = \binom{n}{n_1, \dots, n_s} \sum_{j=1}^s \frac{n_j}{n} d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s} \cdot d_{[\lambda_1]} \cdots d_{[\lambda_s]} \cdot \frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}},$$

for $2 \leq k \leq s$, from which the desired result follows.

Finally, we consider the elements $(\vec{v}; \tau) \in \tilde{G} = G \wr S_n$, which comprise s conjugacy classes: $C_1^{(\vec{v}; \tau)}, C_2^{(\vec{v}; \tau)}, \dots, C_s^{(\vec{v}; \tau)}$, where the indices are chosen so that, for $C_k^{(\vec{v}; \tau)}$, the product (calculated as in the determination of the type matrices in Section 3.3) of the two possible non-identity entries of \vec{v} is in conjugacy class C_k of G . Since there are $\frac{1}{2}n(n-1) \cdot |G| \cdot |C_k|$ elements in the conjugacy class $C_k^{(\vec{v}; \tau)}$ of \tilde{G} , we have $\ell = \frac{1}{2}n(n-1) \cdot |G| \cdot |C_k|$.

Each of the s conjugacy classes splits into s classes in \tilde{H} : in the first class, τ transposes two of the first n_1 elements leaving other elements fixed; in the second class, τ transposes two of the next n_2 elements leaving the other elements fixed; etc. Also, in the first class, the only non-identity entries of \vec{v} are in its first n_1 entries; in the second class, the only non-identity entries of \vec{v} are in its next n_2 entries; etc. Thus $t = s$ with $k_j = \frac{1}{2}n_j(n_j-1) \cdot |G| \cdot |C_k|$ for all $1 \leq j \leq s$.

For a particular conjugacy class C_k of G , let g_k be a representative element. Then the characters of the s classes of elements $(\vec{v}; \tau) \in \tilde{H} = G \wr S_{(n)}$ are

$$d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s} \cdot d_{[\lambda_1]} \cdots d_{[\lambda_s]} \cdot \frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}} \cdot \frac{\chi_{[\lambda_j]}(\tau)}{d_{[\lambda_j]}},$$

for $1 \leq j \leq s$.

Therefore, it follows from Lemma 2.4.3 that the character at $(\vec{v}; \tau)$ of the induced irreducible representation ρ of $G \wr S_n$ is

$$\chi_\rho(\vec{v}, \tau) = \binom{n}{n_1, \dots, n_s} \sum_{j=1}^s \frac{n_j(n_j-1)}{n(n-1)} d_{\rho_1}^{n_1} \cdots d_{\rho_s}^{n_s} \cdot d_{[\lambda_1]} \cdots d_{[\lambda_s]} \cdot \frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}} \cdot \frac{\chi_{[\lambda_j]}(\tau)}{d_{[\lambda_j]}},$$

for $1 \leq k \leq s$, from which the desired result follows. \square

3.6. Analysis of the Independent Shuffles Random Walk. In order to continue our analysis of the independent shuffles random walk introduced in Section 3.1, we must now calculate the Fourier transform of P at each irreducible representation of the complete monomial group $G \wr S_n$.

LEMMA 3.6.1. *Let P be the probability measure on $G \wr S_n$ defined in (3.1.1). For the irreducible representation ρ of $G \wr S_n$ induced by the representation $\rho_{(n)} \otimes \rho_{(\lambda)}$ of $G \wr S_{(n)}$, the Fourier transform is*

$$\hat{P}(\rho) = \left[\frac{n_1}{n^2} + \frac{n_1(n_1-1)}{n^2} r(\lambda_1) \right] I,$$

where $(n) = (n_1, \dots, n_s)$, $(\lambda) = ([\lambda_1], \dots, [\lambda_s])$, and $r(\lambda_1) = \chi_{[\lambda_1]}(\tau)/d_{[\lambda_1]}$ with transposition $\tau \in S_{n_1}$.

PROOF. Recall from Lemma 3.3.1 that P is constant on conjugacy classes. It then follows from Lemma 2.5.1 that $\widehat{P}(\rho) = C \cdot I$, where C is a constant. By applying the results from Corollary 3.3.2 and Lemma 3.5.1, we find that

$$\begin{aligned}
C &= \frac{1}{|G|n}(1)(1) + \sum_{k=2}^s \frac{1}{|G|n^2} (n|C_k|) \sum_{j=1}^s \left(\frac{n_j}{n}\right) \frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}} \\
&+ \sum_{k=1}^s \frac{2}{|G|^2 n^2} \cdot \frac{n(n-1)}{2} \cdot |G| \cdot |C_k| \sum_{j=1}^s \frac{n_j(n_j-1)}{n(n-1)} \cdot \frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}} \cdot r(\lambda_j) \\
&= \frac{1}{|G|n} + \frac{1}{|G|n^2} \sum_{j=1}^s \frac{n_j}{d_{\rho_j}} \sum_{k=2}^s |C_k| \cdot \chi_{\rho_j}(g_k) \\
&+ \frac{1}{|G|n^2} \sum_{j=1}^s \frac{n_j(n_j-1)}{d_{\rho_j}} r(\lambda_j) \sum_{k=1}^s |C_k| \cdot \chi_{\rho_j}(g_k).
\end{aligned}$$

When $j = 1$, since ρ_j is the trivial representation of G with $d_{\rho_j} = 1$, we have $\sum_{k=1}^s |C_k| \cdot \chi_{\rho_j}(g_k) = \sum_{k=1}^s |C_k| = |G|$ and $\sum_{k=2}^s |C_k| \cdot \chi_{\rho_j}(g_k) = |G| - |C_1| = |G| - 1$. When $2 \leq j \leq s$, it follows from Proposition 2.4.4 that $\sum_{k=1}^s |C_k| \cdot \chi_{\rho_j}(g_k) = 0$ and $\sum_{k=2}^s |C_k| \cdot \chi_{\rho_j}(g_k) = -|C_1| \cdot \chi_{\rho_j}(g_1) = -d_{\rho_j}$. Thus

$$\begin{aligned}
C &= \frac{1}{|G|n} + \frac{n_1(|G| - 1)}{|G|n^2} + \frac{(n - n_1)(-1)}{|G|n^2} + \frac{n_1(n_1 - 1)r(\lambda_1)|G|}{|G|n^2} \\
&= \frac{n_1}{n^2} + \frac{n_1(n_1 - 1)}{n^2} r(\lambda_1). \quad \square
\end{aligned}$$

By applying the results from Lemmas 3.5.1 and 3.6.1 to Lemma 2.6.1, we determine all the eigenvalues of the transition matrix \mathbf{P} induced by the probability measure P , together with their multiplicities.

COROLLARY 3.6.2. *Let P be the probability measure on $G \wr S_n$ defined in (3.1.1). Let \mathbf{P} be the transition matrix of the Markov chain induced by the probability measure P . Then, for the irreducible representation ρ of $G \wr S_n$ induced by the representation $\rho_{(n)} \otimes \rho_{(\lambda)}$ of $G \wr S_{(n)}$, there is an eigenvalue π_ρ of \mathbf{P} occurring with algebraic multiplicity*

$$\binom{n}{n_1, \dots, n_s}^2 d_{\rho_1}^{2n_1} \cdots d_{\rho_s}^{2n_s} \cdot d_{[\lambda_1]}^2 \cdots d_{[\lambda_s]}^2$$

such that

$$\pi_\rho = \frac{n_1}{n^2} + \frac{n_1(n_1-1)}{n^2}r(\lambda_1),$$

where $(n) = (n_1, \dots, n_s)$, $(\lambda) = ([\lambda_1], \dots, [\lambda_s])$, and $r(\lambda_1) = \chi_{[\lambda_1]}(\tau)/d_{[\lambda_1]}$ with transposition $\tau \in S_{n_1}$.

We have now established the results necessary to prove Theorem 3.1.3.

PROOF OF THEOREM 3.1.3. By applying the results from Lemmas 3.5.1 and 3.6.1 to the Upper Bound Lemma (2.6.2), we find that

$$\|P^{*k} - U\|_{\text{TV}}^2 \leq \frac{1}{4} (|G|^n n!) \|P^{*k} - U\|_2^2 = \frac{1}{4} \sum_{\rho} d_{\rho}^2 \left[\frac{n_1}{n^2} + \frac{n_1(n_1-1)}{n^2} r(\lambda_1) \right]^{2k},$$

where the sum is taken over all nontrivial irreducible representations of $G \wr S_n$. Thus we have

$$\begin{aligned} \|P^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|P^{*k} - U\|_2^2 \\ &= \frac{1}{4} \sum_{(n)} \sum_{(\lambda)} \binom{n}{n_1, \dots, n_s}^2 d_{\rho_1}^{2n_1} \cdots d_{\rho_s}^{2n_s} \cdot d_{[\lambda_1]}^2 \cdots d_{[\lambda_s]}^2 \left[\frac{n_1}{n^2} + \frac{n_1(n_1-1)}{n^2} r(\lambda_1) \right]^{2k} \\ &= \frac{1}{4} \sum_{n_1=0}^n \binom{n}{n_1}^2 d_{\rho_1}^{2n_1} \sum_{[\lambda_1]} d_{[\lambda_1]}^2 \left[\frac{n_1}{n^2} + \frac{n_1(n_1-1)}{n^2} r(\lambda_1) \right]^{2k} \\ &\quad \times \sum_{(n_2, \dots, n_s)} \binom{n-n_1}{n_2, \dots, n_s}^2 d_{\rho_2}^{2n_2} \cdots d_{\rho_s}^{2n_s} \sum_{([\lambda_2], \dots, [\lambda_s])} d_{[\lambda_2]}^2 \cdots d_{[\lambda_s]}^2. \end{aligned}$$

Consider the direct product $S_{n_2} \times \cdots \times S_{n_s}$. It follows from Proposition 2.3.5 and Lemma 2.3.2 that the sum of the squares of the dimensions of all the irreducible representations of $S_{n_2} \times \cdots \times S_{n_s}$ is

$$\sum_{([\lambda_2], \dots, [\lambda_s])} d_{[\lambda_2]}^2 \cdots d_{[\lambda_s]}^2 = |S_{n_2} \times \cdots \times S_{n_s}| = n_2! \cdots n_s!.$$

Using the multinomial theorem and Lemma 2.3.2, we also have

$$\begin{aligned} \sum_{(n_2, \dots, n_s)} \binom{n-n_1}{n_2, \dots, n_s} d_{\rho_2}^{2n_2} \cdots d_{\rho_s}^{2n_s} &= (d_{\rho_2}^2 + \cdots + d_{\rho_s}^2)^{n-n_1} \\ &= (|G| - d_{\rho_1}^2)^{n-n_1} = (|G| - 1)^{n-n_1}. \end{aligned}$$

Thus

$$\sum_{(n_2, \dots, n_s)} \binom{n - n_1}{n_2, \dots, n_s}^2 d_{\rho_2}^{2n_2} \dots d_{\rho_s}^{2n_s} \sum_{([\lambda_2], \dots, [\lambda_s])} d_{[\lambda_2]}^2 \dots d_{[\lambda_s]}^2 = (|G| - 1)^{n - n_1} (n - n_1)!,$$

which combines with the results above to give

$$\begin{aligned} \|P^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|P^{*k} - U\|_2^2 \\ &= \frac{1}{4} \sum_{n_1=0}^n \binom{n}{n_1} \frac{n!}{n_1!} (|G| - 1)^{n - n_1} \sum_{[\lambda_1]} d_{[\lambda_1]}^2 \left[\frac{n_1}{n^2} + \frac{n_1(n_1 - 1)}{n^2} r(\lambda_1) \right]^{2k}, \end{aligned} \quad (3.6.3)$$

where the inner sum is taken over all partitions $[\lambda_1]$ of n_1 . The trivial representation of $G \wr S_n$, which should not be included in the summations above, occurs when $n_1 = n$ and $[\lambda_1] = [n]$ (which gives the trivial representation of S_n).

For each $1 \leq n_1 \leq n$, it follows from (2.7.2) that we may bound the inner sum above, with $[\lambda_1] = [n_1]$ excluded, by

$$\begin{aligned} \sum_{[\lambda_1]} d_{[\lambda_1]}^2 \left[\frac{n_1}{n^2} + \frac{n_1(n_1 - 1)}{n^2} r(\lambda_1) \right]^{2k} &= \left(\frac{n_1}{n} \right)^{4k} \sum_{[\lambda_1]} d_{[\lambda_1]}^2 \left[\frac{1}{n_1} + \frac{n_1 - 1}{n_1} r(\lambda_1) \right]^{2k} \\ &\leq \left(\frac{n_1}{n} \right)^{4k} 4a^2 e^{-4c} \end{aligned}$$

for a universal constant $a > 0$, when $k \geq \frac{1}{2}n_1 \log n_1 + cn_1$. Since $n \geq n_1$ and $|G| \geq 2$, this is also true when $k \geq \frac{1}{2}n \log n + \frac{1}{4}n \log(|G| - 1) + cn$.

We must also bound the term for the trivial representation $[\lambda_1] = [n_1]$ for $1 \leq n_1 \leq n - 1$. Since in these cases, $d_{[\lambda_1]}^2 = 1$ and $r(\lambda_1) = 1$, we have

$$d_{(\lambda_1)}^2 \left[\frac{n_1}{n^2} + \frac{n_1(n_1 - 1)}{n^2} r(\lambda_1) \right]^{2k} = \left[\frac{n_1}{n^2} + \frac{n_1(n_1 - 1)}{n^2} \right]^{2k} = \left(\frac{n_1}{n} \right)^{4k}.$$

These results lead to the upper bound

$$\begin{aligned} \|P^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|P^{*k} - U\|_2^2 \\ &= \frac{1}{4} \sum_{n_1=0}^n \binom{n}{n_1} \frac{n!}{n_1!} (|G| - 1)^{n - n_1} \sum_{[\lambda_1]} d_{[\lambda_1]}^2 \left[\frac{n_1}{n^2} + \frac{n_1(n_1 - 1)}{n^2} r(\lambda_1) \right]^{2k} \\ &\leq a^2 e^{-4c} \sum_{n_1=1}^n \binom{n}{n_1} \frac{n!}{n_1!} (|G| - 1)^{n - n_1} \left(\frac{n_1}{n} \right)^{4k} \\ &\quad + \frac{1}{4} \sum_{n_1=1}^{n-1} \binom{n}{n_1} \frac{n!}{n_1!} (|G| - 1)^{n - n_1} \left(\frac{n_1}{n} \right)^{4k}. \end{aligned}$$

Now notice that, when $k = \frac{1}{2}n \log n + \frac{1}{4}n \log(|G| - 1) + cn$, then

$$\left(\frac{n_1}{n}\right)^{4k} = \left(\frac{n_1}{n}\right)^{-n[-2\log(n) - \log(|G|-1) - 4c]} = \left[\frac{e^{-4c}}{(|G|-1)n^2}\right]^{-n \log(n_1/n)},$$

which combines with the results above to give

$$\begin{aligned} \|P^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|P^{*k} - U\|_2^2 \\ &\leq a^2 e^{-4c} \sum_{n_1=1}^n \binom{n}{n_1} \frac{n!}{n_1!} (|G| - 1)^{n-n_1} \left[\frac{e^{-4c}}{(|G|-1)n^2}\right]^{-n \log(n_1/n)} \\ &\quad + \frac{1}{4} \sum_{n_1=1}^{n-1} \binom{n}{n_1} \frac{n!}{n_1!} (|G| - 1)^{n-n_1} \left[\frac{e^{-4c}}{(|G|-1)n^2}\right]^{-n \log(n_1/n)}. \end{aligned}$$

If we let $i = n - n_1$, it follows that

$$\begin{aligned} \|P^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|P^{*k} - U\|_2^2 \\ &\leq a^2 e^{-4c} \sum_{i=0}^{n-1} \frac{1}{i!} (e^{-4c})^i + \frac{1}{4} e^{-4c} \sum_{i=0}^{n-2} \frac{1}{(i+1)!} (e^{-4c})^i \\ &\leq a^2 e^{-4c} \exp(e^{-4c}) + \frac{1}{4} e^{-4c} \exp(e^{-4c}). \end{aligned}$$

Since $c > 0$, we have $\exp(e^{-4c}) < e$. Therefore

$$\|P^{*k} - U\|_{\text{TV}}^2 \leq \frac{1}{4} (|G|^n n!) \|P^{*k} - U\|_2^2 \leq \left[\left(a^2 + \frac{1}{4}\right) e\right] e^{-4c},$$

from which the desired result follows. \square

Theorem 3.1.3 shows that $k = \frac{1}{2}n \log n + \frac{1}{4}n \log(|G| - 1) + cn$ steps are sufficient for the (normalized) ℓ^2 distance, and hence also the total variation distance, to become small. A lower bound in the (normalized) ℓ^2 metric can also be derived by examining $n^2(|G| - 1) \left(1 - \frac{1}{n}\right)^{4k}$, which is the dominant contribution to the summation (3.6.3) from the proof of Theorem 3.1.3. This term corresponds to the choice $n_1 = n - 1$ with $[\lambda_1] = [n - 1]$. Notice that $k = \frac{1}{2}n \log n + \frac{1}{4}n \log(|G| - 1) - cn$ steps are necessary for just this term to become small.

That $k = \frac{1}{2}n \log n - cn$ steps are *necessary* for the total variation distance to become small follows directly from Theorem 2.7.3, as follows. Recall that Theorem 2.7.3 shows that $k = \frac{1}{2}n \log n - cn$ steps are necessary for the total variation distance to uniformity to become small for a random walk generated by random transpositions from the symmetric group S_n . This is exactly the random walk on $G \wr S_n$ introduced in Section 3.1, if the vector \vec{x} from $(\vec{x}; \pi) \in G \wr S_n$ is ignored. Thus Theorem 2.7.3 provides a lower bound on the distance to uniformity in the total variation metric, and hence also in the (normalized) ℓ^2 metric.

In the special case of the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, we have $|G| = 2$. Thus the upper bound in Theorem 3.1.3 matches the lower bound derived from Theorem 2.7.3.

That $k = \frac{1}{2}n \log n + cn$ steps are also *sufficient* for the total variation distance to become small is a result of the following.

THEOREM 3.6.4. *Let P and U be the probability measures on the complete monomial group $G \wr S_n$ defined in (3.1.1) and (3.1.2), respectively. Let $k = \frac{1}{2}n \log n + cn$. Then there exists a universal constant $\hat{b} > 0$ such that*

$$\|P^{*k} - U\|_{\text{TV}} \leq \hat{b}e^{-2c} \quad \text{for all } c > 0.$$

PROOF. The probability measure P on $G \wr S_n$ induces probability measures Q on G^n and R on S_n by defining

$$Q(\vec{x}) := \sum_{\pi \in S_n} P(\vec{x}; \pi) \quad \text{and} \quad R(\pi) := \sum_{\vec{x} \in G^n} P(\vec{x}; \pi).$$

Notice that R is the probability measure on S_n defined in (2.1.1).

Recall that P is the one-step distribution for a random walk $\mathbf{W} = (W_0, W_1, W_2, \dots)$ on $G \wr S_n$, as described in Section 2.6. Define stochastic processes $\mathbf{X} = (X_0, X_1, X_2, \dots)$, with state space G^n , and $\mathbf{Y} = (Y_0, Y_1, Y_2, \dots)$, with state space S_n , by setting $W_k =: (X_k; Y_k)$ for $k \geq 0$. It is not hard to see that \mathbf{X} (resp., \mathbf{Y}) is a random walk on G^n (resp., on S_n) with one-step distribution Q (resp., R).

For each $1 \leq i \leq n$, let T_i be the step index k for the random walk \mathbf{W} at which the element $x_i \in \vec{x}$ is first multiplied by a uniformly chosen random element of G , as the result of an element either of the form $(\vec{u}; e)$ or of the form $(\vec{v}; \tau)$. Define $T := \max_{1 \leq i \leq n} T_i$. Thus T is the step at which the last element of \vec{x} is randomized, and

$$\mathbb{P}\{T > k\} \leq \sum_{i=1}^n \mathbb{P}\{T_i > k\} = \sum_{i=1}^n \left(1 - \frac{1}{n}\right)^{2k} \leq ne^{-2k/n}.$$

Notice that

$$\begin{aligned} P^{*k}(\vec{x}; \pi) &\geq \mathbb{P}\{W_k = (\vec{x}; \pi), T \leq k\} \\ &= \frac{1}{|G|^n} \mathbb{P}\{Y_k = \pi, T \leq k\} \\ &= \frac{1}{|G|^n} [R^{*k}(\pi) - \mathbb{P}\{Y_k = \pi, T > k\}]. \end{aligned}$$

Thus

$$\begin{aligned}
\|P^{*k} - U\|_{\text{TV}} &= \max_{A \subseteq G} [U(A) - P^{*k}(A)] \\
&\leq \max_{A \subseteq G} \sum_{(\vec{x}; \pi) \in A} \left[\frac{1}{|G|^n n!} - \frac{1}{|G|^n} R^{*k}(\pi) + \frac{1}{|G|^n} \mathbb{P}\{Y_k = \pi, T > k\} \right] \\
&\leq \max_{A \subseteq G} \sum_{(\vec{x}; \pi) \in A} \left[\frac{1}{|G|^n n!} - \frac{1}{|G|^n} R^{*k}(\pi) \right] + \sum_{(\vec{x}; \pi) \in G \wr S_n} \frac{1}{|G|^n} \mathbb{P}\{Y_k = \pi, T > k\} \\
&= \|R^{*k} - U_{S_n}\|_{\text{TV}} + \mathbb{P}\{T > k\}.
\end{aligned}$$

Let $k = \frac{1}{2}n \log n + cn$. It follows from Theorem 2.1.3 that there exists a universal constant $a > 0$ such that $\|R^{*k} - U_{S_n}\|_{\text{TV}} \leq ae^{-2c}$ for all $c > 0$, where U_{S_n} is the uniform distribution on S_n as defined in (2.1.2). Furthermore, it follows from above that $\mathbb{P}\{T > k\} \leq e^{-2c}$. Therefore

$$\|P^{*k} - U\|_{\text{TV}} \leq (a + 1)e^{-2c},$$

from which the desired result follows. \square

Notice that, for the independent shuffles random walk, the rate of convergence to uniformity for the (normalized) ℓ^2 distance is slightly slower (due to the addition of the term $\frac{1}{4}n \log(|G| - 1)$) than that for the total variation distance. However, if $|G|$ is moderate relative to n , these rates of convergence are “essentially” the same.

The following table summarizes the number of steps (both necessary and sufficient) for the distance (both normalized ℓ^2 and total variation) to uniformity to become small for various special cases of the independent shuffles random walk analyzed in this section.

Random walk on $G \wr S_n$ (with independent randomizations)				
G	metric	nec. or suff.	number of steps	proof
\mathbb{Z}_2	ℓ^2	sufficient	$\frac{1}{2}n \log n$	Thm. 3.1.3
		necessary	$\frac{1}{2}n \log n$	pf. of Thm. 3.1.3
	TV	sufficient	$\frac{1}{2}n \log n$	Thm. 3.1.3
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3
\mathbb{Z}_m	ℓ^2	sufficient	$\frac{1}{2}n \log n + \frac{1}{4}n \log(m-1)$	Thm. 3.1.3
		necessary	$\frac{1}{2}n \log n + \frac{1}{4}n \log(m-1)$	pf. of Thm. 3.1.3
	TV	sufficient	$\frac{1}{2}n \log n$	Thm. 3.6.4
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3
S_m	ℓ^2	sufficient	$\frac{1}{2}n \log n + \frac{1}{4}n \log(m! -1)$	Thm. 3.1.3
		necessary	$\frac{1}{2}n \log n + \frac{1}{4}n \log(m! -1)$	pf. of Thm. 3.1.3
	TV	sufficient	$\frac{1}{2}n \log n$	Thm. 3.6.4
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3
G abelian	ℓ^2	sufficient	$\frac{1}{2}n \log n + \frac{1}{4}n \log(G -1)$	Thm. 3.1.3
		necessary	$\frac{1}{2}n \log n + \frac{1}{4}n \log(G -1)$	pf. of Thm. 3.1.3
	TV	sufficient	$\frac{1}{2}n \log n$	Thm. 3.6.4
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3
G nonabelian	ℓ^2	sufficient	$\frac{1}{2}n \log n + \frac{1}{4}n \log(G -1)$	Thm. 3.1.3
		necessary	$\frac{1}{2}n \log n + \frac{1}{4}n \log(G -1)$	pf. of Thm. 3.1.3
	TV	sufficient	$\frac{1}{2}n \log n$	Thm. 3.6.4
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3

3.7. Analysis of the Paired Shuffles Random Walk. We now describe a slight variant of the independent shuffles random walk introduced in Section 3.1. This will provide a second benchmark random walk, with known rate of convergence, on $G \wr S_n$ for use in the comparison technique. Schoolfield (1998) analyzed a random walk for which comparisons to the independent shuffles and paired shuffles random walks, in the special case of the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, gave different bounds.

Imagine n decks of cards, labeled 1 through n , in sequential order, each with its m cards in sequential order. Independently choose two integers p and q , uniformly from $\{1, 2, \dots, n\}$.

If $p \neq q$, transpose the decks in positions p and q . Then, independently of the choice of p and q and uniformly (i.e., with probability $\frac{1}{|G|} = \frac{1}{m!}$ each), permute the deck terminating in position p by a permutation $\pi \in G = S_m$ and permute the deck terminating in position q by $\pi^{-1} \in G = S_m$. Notice that the only elements of the form $(\vec{v}; \tau)$ that occur in this combination of operations are from the single conjugacy class $C_1^{(\vec{v}; \tau)}$. The probability that an element from any of the other $s - 1$ conjugacy classes occurs now vanishes.

If $p = q$ (which occurs with probability $1/n$), leave the decks in their current positions. Then, again independently and uniformly, permute the deck in position $p = q$ by a permutation in $G = S_m$. The probabilities of the identity and of the elements $(\vec{u}; e)$ are thus unchanged from the independent shuffles random walk.

Again, as in Section 3.1, we will actually examine a random walk on $G \wr S_n$ for any group G , not just the symmetric group S_m . In this more general case, the example above is equivalent to beginning with a vector $(e, \dots, e) \in G^n$, where each $e \in G$ is the identity element. Two elements of this vector are then transposed as the decks of cards were above. The transposed elements of this vector are then multiplied by elements of G as the individual decks were permuted above.

We refer to the process on $G \wr S_n$ described above as the *paired shuffles* random walk, again retaining use of the word “shuffles” even when G is not necessarily S_m . As in Section 3.1, the *paired spins* and *paired flips* random walks are defined in an analogous manner in the special cases of the generalized symmetric group $\mathbb{Z}_m \wr S_n$ and the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, respectively.

The paired shuffles random walk may be modeled formally by a probability measure Q on the complete monomial group $G \wr S_n$. We may thus define the following probability measure

on the set of all elements of $G \wr S_n$:

$$\begin{aligned}
Q(\vec{e}; e) &= \frac{1}{|G|n}, \\
Q(\vec{u}; e) &= \frac{1}{|G|n^2} \text{ where } \vec{u} \neq \vec{e} \in G^n, \\
Q(\vec{v}; \tau) &= \frac{2}{|G|n^2} \text{ where } (\vec{v}; \tau) \in C_1^{(\vec{v}; \tau)}, \\
Q(\vec{x}; \pi) &= 0 \quad \text{otherwise,}
\end{aligned} \tag{3.7.1}$$

where there is only one non-identity entry of $\vec{u} \in G^n$, and where if $\tau \in S_n$ is the transposition $(p \ q)$ then the only possible non-identity entries of $\vec{v} \in G^n$ are in positions p and q and these entries are mutually inverse elements of G . In the special case of the hyperoctahedral group $\mathbb{Z}_2 \wr S_n$, the conjugacy class $C_1^{(\vec{v}; \tau)}$ is the even transpositions.

In order to continue our analysis of the paired shuffles random walk, we must now calculate the Fourier transform of Q at each irreducible representation of the complete monomial group $G \wr S_n$.

LEMMA 3.7.2. *Let Q be the probability measure on $G \wr S_n$ defined in (3.7.1). For the irreducible representation ρ of $G \wr S_n$ induced by the representation $\rho_{(n)} \otimes \rho_{(\lambda)}$ of $G \wr S_{(n)}$, the Fourier transform is*

$$\widehat{Q}(\rho) = \left[\frac{n_1}{n^2} + \sum_{j=1}^s \frac{n_j(n_j - 1)}{n^2} r(\lambda_j) \right] I,$$

where $(n) = (n_1, \dots, n_s)$, $(\lambda) = ([\lambda_1], \dots, [\lambda_s])$, and $r(\lambda_j) = \chi_{[\lambda_j]}(\tau)/d_{[\lambda_j]}$ with transposition $\tau \in S_{n_j}$.

PROOF. Notice that Q is constant on the conjugacy classes of $G \wr S_n$. It then follows from Lemma 2.5.1 that $\widehat{Q}(\rho) = C \cdot I$, where C is a constant. By applying the results from Corollary 3.3.2 and Lemma 3.5.1, we find that

$$\begin{aligned}
C &= \frac{1}{|G|n}(1)(1) + \sum_{k=2}^s \frac{1}{|G|n^2} (n|C_k|) \sum_{j=1}^s \left(\frac{n_j}{n}\right) \frac{\chi_{\rho_j}(g_k)}{d_{\rho_j}} \\
&\quad + \frac{2}{|G|n^2} \cdot \frac{n(n-1)}{2} \cdot |G| \sum_{j=1}^s \frac{n_j(n_j-1)}{n(n-1)} r(\lambda_j) \\
&= \frac{1}{|G|n} + \frac{1}{|G|n^2} \sum_{j=1}^s \frac{n_j}{d_{\rho_j}} \sum_{k=2}^s |C_k| \cdot \chi_{\rho_j}(g_k) + \sum_{j=1}^s \frac{n_j(n_j-1)}{n^2} r(\lambda_j) \\
&= \frac{1}{|G|n} + \frac{n_1(|G|-1)}{|G|n^2} + \frac{(n-n_1)(-1)}{|G|n^2} + \sum_{j=1}^s \frac{n_j(n_j-1)}{n^2} r(\lambda_j) \\
&= \frac{n_1}{n^2} + \sum_{j=1}^s \frac{n_j(n_j-1)}{n^2} r(\lambda_j). \quad \square
\end{aligned}$$

By applying the results from Lemmas 3.5.1 and 3.7.2 to Lemma 2.6.1, we determine all the eigenvalues of the transition matrix \mathbf{Q} induced by the probability measure Q , together with their multiplicities.

COROLLARY 3.7.3. *Let Q be the probability measure on $G \wr S_n$ defined in (3.7.1). Let \mathbf{Q} be the transition matrix of the Markov chain induced by the probability measure Q . Then, for the irreducible representation ρ of $G \wr S_n$ induced by the representation $\rho_{(n)} \otimes \rho_{(\lambda)}$ of $G \wr S_{(n)}$, there is an eigenvalue π_ρ of \mathbf{Q} occurring with algebraic multiplicity*

$$\left(\binom{n}{n_1, \dots, n_s} \right)^2 d_{\rho_1}^{2n_1} \dots d_{\rho_s}^{2n_s} \cdot d_{[\lambda_1]}^2 \dots d_{[\lambda_s]}^2$$

such that

$$\pi_\rho = \frac{n_1}{n^2} + \sum_{j=1}^s \frac{n_j(n_j-1)}{n^2} r(\lambda_j),$$

where $(n) = (n_1, \dots, n_s)$, $(\lambda) = ([\lambda_1], \dots, [\lambda_s])$, and $r(\lambda_j) = \chi_{[\lambda_j]}(\tau)/d_{[\lambda_j]}$ with transposition $\tau \in S_{n_j}$.

The following result establishes an upper bound on both the total variation distance and the ℓ^2 distance between Q^{*k} and the uniform distribution U on $G \wr S_n$. The total variation upper bound is rather poor when G is nonabelian, as shown by Theorem 3.7.6. The quality of the ℓ^2 upper bound will be discussed following the proof of the theorem.

THEOREM 3.7.4. *Let Q and U be the probability measures on the complete monomial group $G \wr S_n$ defined in (3.7.1) and (3.1.2), respectively. Let*

$$k = \max \left\{ n \log n + \frac{1}{2} n \log(|G| - 1) + \frac{1}{2} n \log(s - 1), \frac{1}{2} n \log \delta_n \right\} + 2cn,$$

where $\delta_n := \sum_{j=2}^s d_{\rho_j}^{2n}$ and d_{ρ_j} is the dimension of the irreducible representation ρ_j of G for $2 \leq j \leq s$. Then there exists a universal constant $b > 0$ such that

$$\|Q^{*k} - U\|_{\text{TV}} \leq \frac{1}{2} (|G|^n n!)^{1/2} \|Q^{*k} - U\|_2 \leq b e^{-2c} \quad \text{for all } c > 0.$$

PROOF. By applying the results from Lemmas 3.5.1 and 3.7.2 to the Upper Bound Lemma (2.6.2), we find that

$$\begin{aligned} \|Q^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|Q^{*k} - U\|_2^2 = \frac{1}{4} \sum_{\rho} d_{\rho}^2 \left[\frac{n_1}{n^2} + \sum_{j=1}^s \frac{n_j(n_j - 1)}{n^2} r(\lambda_j) \right]^{2k} \\ &= \frac{1}{4} \sum_{(n)} \sum_{(\lambda)} \binom{n}{n_1, \dots, n_s}^2 d_{\rho_1}^{2n_1} \cdots d_{\rho_s}^{2n_s} \cdot d_{[\lambda_1]}^2 \cdots d_{[\lambda_s]}^2 \left[\frac{n_1}{n^2} + \sum_{j=1}^s \frac{n_j(n_j - 1)}{n^2} r(\lambda_j) \right]^{2k} \end{aligned} \quad (3.7.5)$$

where the sums are taken over all nontrivial irreducible representations of $G \wr S_n$. Notice that

$$\begin{aligned} \left[\frac{n_1}{n^2} + \sum_{j=1}^s \frac{n_j(n_j - 1)}{n^2} r(\lambda_j) \right]^{2k} &= \left\{ \left[\frac{n_1}{n^2} + \frac{n_1(n_1 - 1)}{n^2} r(\lambda_1) \right] + \sum_{j=2}^s \frac{n_j(n_j - 1)}{n^2} r(\lambda_j) \right\}^{2k} \\ &= \left\{ \left(\frac{n_1}{n} \right)^2 \left[\frac{1}{n_1} + \frac{n_1 - 1}{n_1} r(\lambda_1) \right] + \sum_{j=2}^s \left(\frac{n_j}{n} \right)^2 \frac{n_j - 1}{n_j} r(\lambda_j) \right\}^{2k} \\ &\leq \max \left\{ \left(\frac{n_1}{n} \right)^{2k} \left[\frac{1}{n_1} + \frac{n_1 - 1}{n_1} r(\lambda_1) \right]^{2k}, \max_{2 \leq j \leq s} \left(\frac{n_j}{n} \right)^{2k} \left[\frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k} \right\} \\ &\leq \left(\frac{n_1}{n} \right)^{2k} \left[\frac{1}{n_1} + \frac{n_1 - 1}{n_1} r(\lambda_1) \right]^{2k} + \sum_{j=2}^s \left(\frac{n_j}{n} \right)^{2k} \left[\frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k}, \end{aligned}$$

where the first inequality is due to the fact that $\left(\sum_{j=1}^s \alpha_j x_j \right)^{2k} \leq \max_{1 \leq j \leq s} x_j^{2k}$ whenever each $\alpha_j \geq 0$ and $\alpha_1 + \cdots + \alpha_s = 1$. As noted in the proof of Theorem 5 in Section D of Chapter 3 of Diaconis (1988), to every representation (λ_j) there corresponds a conjugate representation (λ'_j) such that $r(\lambda_j) = -r(\lambda'_j)$. So we have

$$\sum_{[\lambda_j]} \left[\frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k} \leq 2 \sum_{[\lambda_j]: r(\lambda_j) \geq 0} \left[\frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k}.$$

Thus

$$\begin{aligned} \sum_{[\lambda_j]} \left[\frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k} &\leq 2 \sum_{[\lambda_j]: r(\lambda_j) \geq 0} \left[\frac{1}{n_j} + \frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k} \\ &\leq 2 \sum_{[\lambda_j]} \left[\frac{1}{n_j} + \frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k}, \end{aligned}$$

which combines with the previous results to give

$$\begin{aligned} \|Q^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|Q^{*k} - U\|_2^2 \\ &\leq \frac{1}{2} \sum_{(n)} \sum_{(\lambda)} \binom{n}{n_1, \dots, n_s}^2 d_{\rho_1}^{2n_1} \dots d_{\rho_s}^{2n_s} \cdot d_{[\lambda_1]}^2 \dots d_{[\lambda_s]}^2 \sum_{j=1}^s \left(\frac{n_j}{n} \right)^{2k} \left[\frac{1}{n_j} + \frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k} \\ &\quad + \frac{1}{4} \sum_{j=2}^s d_{\rho_j}^{2n} \left(\frac{n-1}{n} \right)^{2k}, \end{aligned}$$

where the sum over $(\lambda) = ([\lambda_1], \dots, [\lambda_s])$ is taken over all partitions of (n_1, \dots, n_s) , except that we omit the trivial partitions $[\lambda_j] = [n]$ for $1 \leq j \leq s$. The final term reintroduces the appropriate terms for $[\lambda_j] = [n]$ for $2 \leq j \leq s$.

Continuing as in the proof of Theorem 3.1.3, this may be simplified to

$$\begin{aligned} \|Q^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|Q^{*k} - U\|_2^2 \\ &\leq \frac{1}{2} \sum_{j=1}^s \sum_{n_j=0}^n \binom{n}{n_j} \frac{n!}{n_j!} (|G| - d_{\rho_j}^2)^{n-n_j} \sum_{[\lambda_j] \vdash n_j} d_{[\lambda_j]}^2 \left(\frac{n_j}{n} \right)^{2k} \left[\frac{1}{n_j} + \frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k} \\ &\quad + \frac{1}{4} \sum_{j=2}^s d_{\rho_j}^{2n} \left(\frac{n-1}{n} \right)^{2k} \\ &\leq \frac{1}{2} s \sum_{n_j=0}^n \binom{n}{n_j} \frac{n!}{n_j!} (|G| - 1)^{n-n_j} \sum_{[\lambda_j] \vdash n_j} d_{[\lambda_j]}^2 \left(\frac{n_j}{n} \right)^{2k} \left[\frac{1}{n_j} + \frac{n_j - 1}{n_j} r(\lambda_j) \right]^{2k} \\ &\quad + \frac{1}{4} \left(\frac{n-1}{n} \right)^{2k} \delta_n \end{aligned}$$

where, for $1 \leq j \leq s$, the sum $\sum_{[\lambda_j] \vdash n_j}$ is taken over all partitions $[\lambda_j]$ of n_j , with $[\lambda_j] = [n]$ excluded when $n_j = n$, and where $\delta_n = \sum_{j=2}^s d_{\rho_j}^{2n}$.

Recall from the proof of Theorem 3.1.3 that, when $k \geq \frac{1}{2}n \log n + \frac{1}{4}n \log(|G| - 1) + c'n$, we may bound the inner sum above using

$$\sum_{[\lambda_1]} d_{[\lambda_1]}^2 \left(\frac{n_1}{n} \right)^{2k} \left[\frac{1}{n_1} + \frac{n_1 - 1}{n_1} r(\lambda_1) \right]^{2k} \leq 4a^2 e^{-4c'} \left(\frac{n_1}{n} \right)^{2k} + \left(\frac{n_1}{n} \right)^{2k}.$$

for $1 \leq n_1 \leq n - 1$, and using

$$\sum_{[\lambda_1]} d_{[\lambda_1]}^2 \left(\frac{n_1}{n} \right)^{2k} \left[\frac{1}{n_1} + \frac{n_1 - 1}{n_1} r(\lambda_1) \right]^{2k} \leq 4a^2 e^{-4c'}$$

for $n_1 = n$. So this is also true when $k \geq n \log n + \frac{1}{2}n \log(|G| - 1) + \frac{1}{2}n \log(s - 1) + 2cn$, in which case we may choose $c' = \frac{1}{2} \log n + \frac{1}{4} \log(|G| - 1) + \frac{1}{2} \log(s - 1) + 2c$. Thus

$$e^{-4c'} = \frac{e^{-8c}}{(|G| - 1)(s - 1)n^2}.$$

Now notice that when $k \geq n \log n + \frac{1}{2}n \log(|G| - 1) + \frac{1}{2}n \log(s - 1) + 2cn$,

$$\left(\frac{n_1}{n} \right)^{2k} \leq \left[\frac{e^{-4c}}{(|G| - 1)(s - 1)n^2} \right]^{-n \log(n_1/n)}.$$

These results lead to the upper bound

$$\begin{aligned} \|Q^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|Q^{*k} - U\|_2^2 \\ &\leq 2sa^2 \left[\frac{e^{-8c}}{(|G| - 1)(s - 1)n^2} \right] \sum_{n_1=1}^n \binom{n}{n_1} \frac{n!}{n_1!} (|G| - 1)^{n-n_1} \left[\frac{e^{-4c}}{(|G| - 1)(s - 1)n^2} \right]^{-n \log(n_1/n)} \\ &\quad + \frac{1}{2} s \sum_{n_1=1}^{n-1} \binom{n}{n_1} \frac{n!}{n_1!} (|G| - 1)^{n-n_1} \left[\frac{e^{-4c}}{(|G| - 1)(s - 1)n^2} \right]^{-n \log(n_1/n)} \\ &\quad + \frac{1}{4} \delta_n \left[\frac{e^{-4c}}{(|G| - 1)(s - 1)n^2} \right]^{-n \log(1 - \frac{1}{n})}. \end{aligned}$$

Continuing as in the proof of Theorem 3.1.3, we find that

$$\begin{aligned} \|Q^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|Q^{*k} - U\|_2^2 \\ &\leq 2a^2 \left[\frac{s}{(|G|-1)(s-1)n^2} \right] e^{-8c} \exp(e^{-4c}) + \frac{1}{2} \left[\frac{s}{s-1} \right] e^{-4c} \exp(e^{-4c}) \\ &\quad + \frac{1}{4} \delta_n \left[\frac{e^{-4c}}{(|G|-1)(s-1)n^2} \right]. \end{aligned}$$

Since $c > 0$, we have $\exp(e^{-4c}) < e$. It then follows that

$$\begin{aligned} \|Q^{*k} - U\|_{\text{TV}}^2 &\leq \frac{1}{4} (|G|^n n!) \|Q^{*k} - U\|_2^2 \\ &\leq [(4a^2 + 1)e] e^{-4c} + \left[\frac{\delta_n}{4(|G| - 1)(s - 1)n^2} \right] e^{-4c}. \end{aligned}$$

Recall that when G is abelian, $\delta_n = \sum_{j=2}^s d_{\rho_j}^{2n} = |G| - 1$ for all n ; in this case, the proof is complete.

When G is nonabelian, δ_n increases exponentially with n . We must thus reexamine the term $\frac{1}{4} \delta_n \left(\frac{n-1}{n} \right)^{2k}$. Notice that when $k \geq \frac{1}{2}n \log \delta_n + 2cn$,

$$\frac{1}{4} \delta_n \left(\frac{n-1}{n} \right)^{2k} \leq \frac{1}{4} \delta_n e^{-2k/n} \leq \frac{1}{4} e^{-4c}.$$

Therefore, choosing $k = \max \left\{ n \log n + \frac{1}{2}n \log(|G| - 1) + \frac{1}{2}n \log(s - 1), \frac{1}{2}n \log \delta_n \right\} + 2cn$ completes the proof. \square

Theorem 3.7.4 shows that $k = \max \left\{ n \log n + \frac{1}{2}n \log(|G| - 1) + \frac{1}{2}n \log(s - 1), \frac{1}{2}n \log \delta_n \right\} + 2cn$ steps are sufficient for the (normalized) ℓ^2 distance, and hence the total variation distance, to become small. When $n \log n + \frac{1}{2}n \log(|G| - 1) + \frac{1}{2}n \log(s - 1) \geq \frac{1}{2}n \log \delta_n$ (which is always the case when G is abelian), a lower bound in the (normalized) ℓ^2 metric can also be derived by examining

$$n^2 \left(1 - \frac{1}{n} \right)^{4k} \sum_{j=2}^s d_{\rho_j}^2 = n^2 (|G| - 1) \left(1 - \frac{1}{n} \right)^{4k},$$

which, in this case, is the dominant contribution to the summation (3.7.5) from the proof of Theorem 3.7.4. This comes from summing (over $2 \leq j \leq s$) the terms corresponding to the choice $n_1 = n - 1$ with $[\lambda_1] = [n - 1]$ and $n_j = 1$ with $[\lambda_j] = [1]$. Notice that $k = \frac{1}{2}n \log n + \frac{1}{4}n \log(|G| - 1) - cn$ steps are necessary for just this term to become small.

Notice that, in this case, since $s \leq |G|$, our upper $[n \log n + \frac{1}{2}n \log(|G| - 1) + \frac{1}{2}n \log(s - 1)]$ and lower $[\frac{1}{2}n \log n + \frac{1}{4}n \log(|G| - 1)]$ bounds on the number of steps required for the (normalized) ℓ^2 distance to become small differ by at most a constant factor, but we have not been able to close this gap. However, this gap will pose no problems in the implementation of the comparison technique, since its results are accurate only up to a constant factor. Nonetheless, no such gap exists for total variation distance, as will be shown later.

When $n \log n + \frac{1}{2}n \log(|G| - 1) + \frac{1}{2}n \log(s - 1) \leq \frac{1}{2}n \log \delta_n$, a matching lower bound in the (normalized) ℓ^2 metric can also be derived by examining

$$\left(\sum_{j=2}^s d_{\rho_j}^{2n} \right) \left(\frac{n-1}{n} \right)^{2k} = \delta_n \left(1 - \frac{1}{n} \right)^{2k},$$

which, in this case, is the dominant contribution to the summation (3.7.5) from the proof of Theorem 3.7.4. This comes from summing (over $2 \leq j \leq s$) the terms corresponding to the choice $n_j = n$ with $[\lambda_j] = [n]$. Notice that $k = \frac{1}{2}n \log \delta_n - cn$ steps are necessary for just this term to become small.

For fixed nonabelian G , notice that $d_{\max}^{2n} \leq \delta_n \leq (s-1)d_{\max}^{2n}$, where $d_{\max} := \max_{2 \leq j \leq s} d_{\rho_j}$. So $2n \log d_{\max} \leq \log \delta_n \leq 2n \log d_{\max} + \log(s-1)$. Thus, in this case, we have thereby determined that very nearly $n^2 \log d_{\max}$ steps are necessary and sufficient to make (normalized) ℓ^2 distance small for a fixed nonabelian group G . How large is d_{\max} ? Since $\sum_{j=2}^s d_{\rho_j}^2 = |G| - 1$, somewhat crude bounds are

$$\left(\frac{|G| - 1}{s - 1} \right)^{1/2} \leq d_{\max} \leq (|G| - 1)^{1/2}.$$

For $G = S_m$ with $m \geq 3$, for example, these bounds are sufficient to show $\log d_{\max}(m) = \frac{1}{2}m \log m - \frac{1}{2}m - O(m^{1/2})$ as $m \rightarrow \infty$, since $|G| = m!$ and $s = p(m) \sim \frac{1}{4m\sqrt{3}} \exp \left\{ \pi \sqrt{\frac{2m}{3}} \right\}$, where the asymptotic formula for the partition function $p(\cdot)$ is due to Hardy and Ramanujan (1918) (see, e.g., Hall (1986), Section 4.2).

That $k = \frac{1}{2}n \log n - cn$ steps are *necessary* for total variation distance to become small again follows directly from Theorem 2.7.3, exactly as in Section 3.6. That $k = \frac{1}{2}n \log n + cn$ steps are also *sufficient* (at least in continuous time) for total variation distance to become small is the result of Theorem 3.7.6, which follows.

All of the random walks studied thus far have been discrete-time random walks. We now introduce the continuous-time analogue of a discrete time random walk. Changing from discrete to continuous time will be advantageous in the proof of Theorem 3.7.6.

Suppose that P is a probability measure defined on a finite group G . The *continuized* chain corresponding to P is the continuous-time Markov chain on G started at the identity e with transition rates

$$p(g, h) = P(hg^{-1})$$

for $g, h \in G$ with $g \neq h$. We denote the distribution of the chain at time t by P_t , which is given by

$$P_t(g) := \sum_{k=0}^{\infty} e^{-t} \frac{t^k}{k!} P^{*k}(g) \quad \text{for } g \in G.$$

The following result shows that time $t_n = \frac{1}{2}n \log n + cn$ is sufficient, as $n \rightarrow \infty$, for the total variation distance to become small in the (continuous-time analogue of the) paired shuffles random walk. The result is established only in the limit because the proof relies on classical random graph results known only (at least to us) in the limit.

THEOREM 3.7.6. *Let Q and U be the probability measures on the complete monomial group $G \wr S_n$ defined in (3.7.1) and (3.1.2), respectively. Let Q_t be the distribution at time t of the continuized chain corresponding to Q . Let $t_n = \frac{1}{2}n \log n + cn$. Then there exists a universal constant $\hat{b} > 0$ such that*

$$\limsup_{n \rightarrow \infty} \|Q_{t_n} - U\|_{\text{TV}} \leq \hat{b}e^{-c} \quad \text{for all } c > 0.$$

PROOF. The probability measure Q on $G \wr S_n$ induces a probability measure R on S_n by defining

$$R(\pi) := \sum_{\vec{x} \in G^n} Q(\vec{x}; \pi).$$

Notice that R is the probability measure on S_n defined in (2.1.1).

In order for the paired shuffles continuized chain to achieve randomness, not only must π in $(\vec{x}; \pi)$ be a random permutation of S_n , but also must the entries of $\vec{x} \in G^n$ be (uniformly) random elements of G . Recall that the values of the entries in positions p and q of \vec{x} are multiplied (on the left) by mutually inverse elements of G when \vec{x} is multiplied (on the left) by $(\vec{v}; \tau) \in C_1^{(\vec{v}; \tau)}$ with $\tau = (p \ q)$. In order to determine the amount of time needed to randomize the entries of \vec{x} , begin with n labeled vertices. At each time that an element $(\vec{v}; (p \ q)) \in C_1^{(\vec{v}; \tau)}$ is generated by Q , consider an edge to be generated between positions p and q in a graph Γ . Let T be the time at which the graph Γ becomes connected.

Let T^* be the first time $t > T$ at which Q generates an element either of the form $(\vec{u}; e)$ or $(\vec{e}; e)$. So we may suppose that at time T^* , an element $x_i \in \vec{x}$ is multiplied (on the left) by a uniformly chosen random element $g \in G$. Thus x_i is now randomized. Since $T^* > T$, it follows that there exist elements $x_{j_1}, \dots, x_{j_m} \in \vec{x}$, with $j_1, \dots, j_m \neq i$ and $m \geq 1$, such that, for $1 \leq \ell \leq m$, at some time $t_{j_\ell} \leq T$, the entry x_{j_ℓ} was multiplied (on the left) by a uniformly chosen random element $h_\ell \in G$ and x_i was multiplied by $h_\ell^{-1} \in G$. But since x_i is now randomized, it follows that x_{j_1}, \dots, x_{j_m} are also. By then considering the elements “paired” with x_{j_1}, \dots, x_{j_m} , and so forth, this argument continues on to show that at time T^* *all* of the elements of \vec{x} are randomized.

Since elements either of the form $(\vec{u}; e)$ or $(\vec{e}; e)$ are generated by Q at exponential rate $\lambda = 1/n$, we have

$$\mathbb{P}\{T^* - T \leq s\} = 1 - e^{-s/n}.$$

It then follows from the independence of T and $T^* - T$ that

$$\begin{aligned} \mathbb{P}\{T^* > t\} &= \int_{s=0}^t \frac{1}{n} e^{-s/n} \mathbb{P}\{T > t - s\} ds \\ &= \int_{u=0}^{\infty} e^{-u} \mathbb{P}\{T > t - un\} \mathbb{I}\{u \leq t/n\} du. \end{aligned}$$

Since each element of the form $(\vec{x}; (p\ q)) \in C_1^{(\vec{x}; \tau)}$, which transposes a particular pair of entries $\{p, q\}$, is generated by Q at exponential rate $\lambda = 2/n^2$, the indicator (call it $I_{\{p,q\}}(t)$) of the presence of any given edge $\{p, q\}$ in Γ at time t has expectation

$$1 - e^{-2t/n^2}.$$

Moreover (and this is the advantage of working in continuous time), the stochastic processes $I_{\{p,q\}}(\cdot)$ are mutually independent. Let $t \equiv t_n = \frac{1}{2}n \log n + cn$. Then

$$1 - e^{-2t/n^2} = 1 - \exp\{-n^{-1}(\log n + 2c)\} \sim n^{-1}(\log n + 2c) \quad \text{as } n \rightarrow \infty$$

for fixed $c \in \mathbb{R}$. It then follows from a classical random graph result of Erdős and Rényi (1959) (see, e.g., Graham et al (1995), Chapter 6, Section 5), that

$$\mathbb{P}\{T > t - un\} = \mathbb{P}\left\{T > \frac{1}{2}n \log n + (c - u)n\right\} \longrightarrow 1 - \exp\{-e^{-2(c-u)}\} \quad \text{as } n \rightarrow \infty$$

for fixed $c, u \in \mathbb{R}$. Thus, by the bounded convergence theorem,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}\{T^* > t\} &= \int_{u=0}^{\infty} e^{-u} [1 - \exp\{-e^{-2(c-u)}\}] du \\ &\leq \int_{u=0}^c e^{-u} e^{-2(c-u)} du + \int_{u=c}^{\infty} e^{-u} du \\ &= e^{-c} - e^{-2c} + e^{-c} \leq 2e^{-c} \end{aligned}$$

for fixed $c \in \mathbb{R}$.

It follows exactly as in the proof of Theorem 3.6.3 that

$$\|Q_t - U\|_{\text{TV}} \leq \|R_t - U_{S_n}\|_{\text{TV}} + \mathbb{P}\{T^* > t\}$$

for every $t \geq 0$. With $t_n = \frac{1}{2}n \log n + cn$, a continuous time analogue of Theorem 2.1.3 (which we have confirmed) asserts that there exists a universal constant $a' > 0$ such that

$$\|R_t - U_{S_n}\|_{\text{TV}} \leq a' e^{-2c} \quad \text{for all } n \geq 1 \text{ and all } c > 0,$$

where U_{S_n} is the uniform distribution on S_n defined by (2.1.2). Therefore,

$$\limsup_{n \rightarrow \infty} \|Q_{t_n} - U\|_{\text{TV}} \leq (a' + 2) e^{-c},$$

from which the desired result follows. \square

The following table summarizes the number of steps (both necessary and sufficient) for the distance (both normalized ℓ^2 and total variation) to uniformity to become small for various special cases of the paired shuffles random walk analyzed in this section.

Random walk on $G \wr S_n$ (with paired randomizations)				
G	metric	nec. or suff.	number of steps	proof
\mathbb{Z}_2	ℓ^2	sufficient	$n \log n$	Thm. 3.7.4
		necessary	$\frac{1}{2}n \log n$	pf. of Thm. 3.7.4
	TV	sufficient	$\frac{1}{2}n \log n \quad (n \rightarrow \infty)$	Thm. 3.7.6
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3
\mathbb{Z}_m	ℓ^2	sufficient	$n \log n + n \log(m-1)$	Thm. 3.7.4
		necessary	$\frac{1}{2}n \log n + \frac{1}{4}n \log(m-1)$	pf. of Thm. 3.7.4
	TV	sufficient	$\frac{1}{2}n \log n \quad (n \rightarrow \infty)$	Thm. 3.7.6
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3
S_m	ℓ^2	sufficient	$\max \left\{ \frac{1}{2}n \log \delta_n, n \log n + \frac{1}{2}n \log(m! - 1) + \frac{1}{2}n \log(p(m) - 1) \right\}$	Thm. 3.7.4
		necessary	$\max \left\{ \frac{1}{2}n \log \delta_n, \frac{1}{2}n \log n + \frac{1}{4}n \log(m! - 1) \right\}$	pf. of Thm. 3.7.4
	TV	sufficient	$\frac{1}{2}n \log n \quad (n \rightarrow \infty)$	Thm. 3.7.6
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3
G abelian	ℓ^2	sufficient	$n \log n + n \log(G - 1)$	Thm. 3.7.4
		necessary	$\frac{1}{2}n \log n + \frac{1}{4}n \log(G - 1)$	pf. of Thm. 3.7.4
	TV	sufficient	$\frac{1}{2}n \log n \quad (n \rightarrow \infty)$	Thm. 3.7.6
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3
G nonabelian	ℓ^2	sufficient	$\max \left\{ \frac{1}{2}n \log \delta_n, n \log n + \frac{1}{2}n \log(G - 1) + \frac{1}{2}n \log(s - 1) \right\}$	Thm. 3.7.4
		necessary	$\max \left\{ \frac{1}{2}n \log \delta_n, \frac{1}{2}n \log n + \frac{1}{4}n \log(G - 1) \right\}$	pf. of Thm. 3.7.4
	TV	sufficient	$\frac{1}{2}n \log n \quad (n \rightarrow \infty)$	Thm. 3.7.6
		necessary	$\frac{1}{2}n \log n$	Thm. 2.7.3

Acknowledgments. This paper formed a portion of the author's Ph.D. dissertation in the Department of Mathematical Sciences at the Johns Hopkins University. The author wishes to thank his advisor Jim Fill, whose assistance was invaluable, particularly in the proof of Theorem 3.7.6. The author also wishes to thank Persi Diaconis for initially suggesting that he analyze a random walk on the hyperoctahedral group. It is from that initial challenge that this paper has evolved.

REFERENCES

- ALPERIN, J. and BELL, R. (1995). *Groups and Representations*. Graduate Texts in Mathematics **162**. Springer-Verlag, New York.
- DIACONIS, P. (1988). *Group Representations in Probability and Statistics*. Institute of Mathematical Statistics, Hayward, CA.
- DIACONIS, P. and SALOFF-COSTE, L. (1993). Comparison techniques for random walk on finite groups. *Ann. Probab.* **21** 2131–2156.
- DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179.
- ERDÖS, P. and RÉNYI, A. (1959). On random graphs I. *Publ. Math. Debrecen* **6** 290–297.
- GRAHAM, R., GRÖTSCHEL, M., and LOVÁSZ, L. (1995). *Handbook of Combinatorics*, Vol. I. Elsevier, Amsterdam.
- HALL, M. (1986). *Combinatorial Theory*, 2nd ed. John Wiley & Sons, New York.
- HARDY, G. and RAMANUJAN, S. (1918). Asymptotic formulae in combinatorial analysis. *Proc. London Math. Soc.* **17** 75–115.
- JAMES, G. and KERBER, A. (1981). *The Representation Theory of the Symmetric Group*. Encyclopedia of Mathematics and its Applications **16**. Addison-Wesley, Reading, MA.
- SAGAN, B. (1991). *The Symmetric Group*. Wadsworth and Brooks/Cole, Pacific Grove, CA.
- SCHOOLFIELD, C. (1998). Random walks on wreath products of groups and Markov chains on related homogeneous spaces. Ph.D. dissertation, Dept. of Mathematical Sciences, The Johns Hopkins University.
- SERRE, J.-P. (1977). *Linear Representations of Finite Groups*. Graduate Texts in Mathematics **42**. Springer-Verlag, New York.
- SIMON, B. (1996). *Representations of Finite and Compact Groups*. Graduate Studies in Mathematics **10**. American Mathematical Society, Providence, RI.

CLYDE H. SCHOOLFIELD, JR.
 DEPARTMENT OF STATISTICS
 HARVARD UNIVERSITY
 ONE OXFORD STREET
 CAMBRIDGE, MASSACHUSETTS 02138
 E-MAIL: clyde@stat.harvard.edu